

UNIVERSIDAD DE CÓRDOBA



Programa de Doctorado interuniversitario Dinámica de Flujos
Biogeoquímicos y su Aplicación

APLICACIONES HIDROLÓGICAS DEL ANÁLISIS MULTIFRACTAL
DE DATOS DE PRECIPITACIÓN

HYDROLOGICAL APPLICATIONS OF MULTIFRACTAL ANALYSIS OF
RAINFALL DATA

Directora: Amanda Penélope García Marín

Autor de la tesis: María Teresa Medina Cobo

Córdoba, Noviembre de 2018

TITULO: *APLICACIONES HIDROLÓGICAS DEL ANÁLISIS MULTIFRACTAL DE
DATOS DE PRECIPITACIÓN*

AUTOR: *María Teresa Medina Cobo*

© Edita: UCOPress. 2019
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

[https://www.uco.es/ucopress/index.php/es/
ucopress@uco.es](https://www.uco.es/ucopress/index.php/es/ucopress@uco.es)

UNIVERSIDAD DE CÓRDOBA

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS AGRÓNOMOS Y
DE MONTES

DEPARTAMENTO DE INGENIERÍA RURAL

**APLICACIONES HIDROLÓGICAS DEL ANÁLISIS
MULTIFRACTAL DE DATOS DE PRECIPITACIÓN**

Tesis Doctoral presentada por María Teresa Medina Cobo, en satisfacción de los requisitos necesarios para optar al grado de Doctor Ingeniero de Montes, y dirigida por la Dra. Amanda Penélope García Marín, de la Universidad de Córdoba.

La Directora:

La doctoranda:



Fdo.: Amanda Penélope García Marín



Fdo.: María Teresa Medina Cobo

Córdoba, Noviembre de 2018



TÍTULO DE LA TESIS:

APLICACIONES HIDROLÓGICAS DEL ANÁLISIS MULTIFRACTAL DE DATOS DE PRECIPITACIÓN

DOCTORANDO/A:

MARÍA TERESA MEDINA COBO

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

Uno de los procesos más importantes en hidrología es la precipitación que exhibe una alta variabilidad tanto en el tiempo como en el espacio. El modelado y la predicción de la naturaleza variable de la precipitación es uno de los objetivos fundamentales de muchas investigaciones llevadas a cabo en los últimos años. Su alta variabilidad ha inducido al estudio de sus diferentes escalas de forma independiente, lo que desemboca en un uso restringido de los modelos hidrológicos. Las teorías de escala se basan en que un proceso exhibe un mismo comportamiento a diferentes escalas de estudio. La teoría de los fractales y su posterior evolución hacia la teoría de los multifractales, estudia matemáticamente esta invarianza de escala.

Inicialmente se propuso como objetivo de esta Tesis la caracterización multifractal de series de datos históricas de precipitación disponibles en diferentes localidades y el uso de los resultados obtenidos en diversas aplicaciones hidrológicas, destacando la agrupación de estaciones dentro del Análisis Regional de precipitación, y la determinación del tiempo mínimo necesario para discriminar eventos de precipitación en un determinado lugar. Ambas aplicaciones se han llevado a cabo de forma exitosa, como queda patente tanto en la Tesis doctoral como en las tres publicaciones derivadas de la misma:

1. García-Marín, A. P., Estévez, J., Medina-Cobo, M. T. y Ayuso-Muñoz, J. L. 2015. Delimiting homogeneous regions using the multifractal properties of validated rainfall data series. *Journal of Hydrology*, 529,106-119.
2. Medina-Cobo, M. T., García-Marín, A. P., Estévez, J. y Ayuso-Muñoz, J. L. 2016. The identification of an appropriate Minimum Inter-event Time (MIT) based on multifractal characterization of rainfall data series. *Hydrological Processes*, 30, 3507–3517.
3. Medina-Cobo, M.T., García-Marín, A. P., Estévez, J, Jiménez-Hornero, F.J., Ayuso-Muñoz, J.L. 2017. Obtaining Homogeneous Regions by Determining the Generalized Fractal Dimensions of Validated Daily Rainfall Data Sets. *Water resources management*, 31, 2333-2348.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 13 de noviembre de 2018

Firma del director



Fdo.: Amanda Penélope García Marín

A mi familia....

Agradecimientos

Al finalizar un trabajo tan complejo como es el desarrollo de una Tesis Doctoral parece inevitable pensar en tu propio mérito. Sin embargo, el análisis objetivo, el pensamiento y el recuerdo de todas y cada una de las etapas te muestran inmediatamente que hubiera sido completamente imposible sin la participación de aquellas personas que han facilitado las cosas para que este trabajo llegara a buen término.

Es por ello que agradezco en primer lugar, de manera especial y sincera, a mi Directora de Tesis Amanda García Marín el aceptarme para realizar este trabajo bajo su dirección. Su apoyo incondicional, su gran ayuda y su confianza en mi trabajo han sido un aporte primordial para mí, ya no sólo en el desarrollo de este trabajo, sino en mi propia formación como investigadora. Gracias por todo Amanda, de corazón.

Y, por supuesto, el agradecimiento más profundo y sentido para mi familia, por su gran apoyo en este camino y, en especial, a Luis y Laura Medina, dos piezas fundamentales en el puzzle que forma mi vida. Luis, gracias por tu apoyo, fuerza, paciencia y generosidad y, sobre todo, gracias por aportarme tiempo para terminar este duro camino. Laura, aunque aún eres muy pequeña para entender todo esto, algún día te lo explicaré y te agradeceré la fuerza que me has aportado.

Luis y Laura, esto es por vosotros.

ÍNDICE

RESUMEN	ix
ABSTRACT	x
 1. INTRODUCCIÓN	 1
 2. FUNDAMENTOS TEÓRICOS.....	 5
2.1. La invarianza de escala: fractales y multifractales.....	5
2.2. Caracterización de procesos multifractales.	8
2.2.1. El formalismo multifractal de turbulencia: el escalado de momentos	8
2.2.2. El formalismo multifractal del “Strange Attractor”.	11
2.3. La teoría de la criticalidad autoorganizada	14
 3. DATOS DE LLUVIA	 17
3.1. Zonas de estudio.....	17
3.1.1. Estaciones seleccionadas.....	20
3.2. Control de Calidad.....	29
3.2.1. Métodos	29

3.2.1.1. <i>Test de Rangos o Intervalos</i>	29
3.2.1.2. <i>Tests basados en decisiones estadísticas</i>	30
3.2.1.3. <i>Tests de Persistencia</i>	31
3.2.1.4. <i>Test de Consistencia Espacial, SRT</i>	32
3.2.1.5. <i>Test de Consistencia espacial SRT Modificado</i>	33
3.2.1.6. <i>Inspección Manual</i>	33
3.2.2. Resultados	34
3.2.2.1. <i>Resultados de validación en las estaciones de la provincia de Málaga (España)</i>	34
3.2.2.2. <i>Resultados de validación en las estaciones de las regiones del Maule y Bio-Bio (Chile)</i>	42
4. APLICACIÓN Y RESULTADOS	43
4.1. Formación de Regiones Homogéneas en base a parámetros multifractales	43
4.1.1. Análisis Regional de Frecuencias.....	43
4.1.1.1 <i>Método del Índice de Avenida</i>	44
4.1.1.2. <i>Momentos Lineales</i>	46
4.1.1.3. <i>Pasos a seguir en el Análisis Regional de Frecuencias</i>	48
4.1.1.4. <i>Discordancia</i>	49
4.1.1.5. <i>Homogeneidad</i>	51
4.1.1.6. <i>Ajuste a una función de distribución de frecuencias</i>	51
4.1.1.7. <i>Métodos para formar regiones</i>	54
4.1.2. Formación de regiones homogéneas de precipitación máxima diaria anual en base a los valores de los parámetros multifractales q_{crit} (momento crítico) y γ_{max}	56

4.1.2.1. Conclusiones.	71
4.1.3. Formación de regiones homogéneas de precipitación máxima diaria anual en base a los valores de los parámetros γ_{\max} y $K(0)$	72
4.1.3.1. Conclusiones.	83
4.1.4 Obtención de regiones homogéneas de precipitación a partir de los valores de la dimensión fractal generalizada	84
4.1.4.1. Análisis multifractal: aplicación del formalismo del StrangeAtractor	84
4.1.4.2. Formación de regiones homogéneas.....	88
4.1.4.3. Conclusiones	93
4.2. Identificación del tiempo mínimo entre eventos de precipitación mediante análisis multifractal.	93
4.2.1. Definición y cálculo de eventos de lluvia	95
4.2.2. Resultados de análisis multifractal y criticalidad autoorganizada	97
4.2.3. Conclusiones	103
5. CONCLUSIONES GENERALES Y LÍNEAS FUTURAS.....	105
6. REFERENCIAS.....	107
7. PUBLICACIONES DERIVADAS DE LA TESIS DOCTORAL	117

INDICE DE FIGURAS

Figura 2.1.	Representación esquemática de los diferentes parámetros del espectro multifractal (Figura obtenida de Krzyszczak J., et al. 2017)	14
Figura 3.1.	Mapa de situación de la provincia de Málaga en España	18
Figura 3.2.	Región del Maule (Chile).....	19
Figura 3.3.	Región del Bio Bio (Chile)	20
Figure 3.4.	Distribución espacial de las estaciones consideradas en la provincia de Málaga.....	24
Figura 3.5.	Distribución espacial de las estaciones consideradas en la región del Maule (Chile).....	27
Figura 3.6.	Fracción marcada de datos en la estación Agujero. Test de Rango Dinámico.....	37
Figura 3.7	Fracción marcada de datos en la estación Alozaina. Test de Persistencia	38
Figura 4.1.	Representación log-log de los momentos q-ésimos. Estación Agujero...61	
Figura 4.2.	Función exponente escaladora de momentos. Estación Agujero.....	65
Figura 4.3.	Representación log-log de los momentos q-ésimos. Estación Andalien. 76	
Figura 4.4.	Función exponente escaladora de momentos. Estación Andalien.	77
Figura 4.5.	Valores de la dimensión fractal generalizada para 4 estaciones de la región del Maule.	85
Figura 4.6.	Espectro multifractal $f(\alpha)$ para las estaciones de Agua Fría, Bullileo, Liguay y San Rafael.....	88
Figura 4.7.	Regiones homogéneas obtenidas para la región del Maule (Chile).....	92
Figura 4.8.	Definición de evento de lluvia	95
Figura 4.9.	Diagrama de cajas y bigotes	97

Figura 4.10.	Funciones $K(q)$ empíricas y teóricas para Bullileo y Málaga.	98
Figura 4. 11.	Ajustes obtenidos con la función de distribución empírica para los tamaños de los eventos de lluvia en Málaga con MTE entre 1 y 6 horas	101
Figura 4.12.	Ajustes obtenidos con la función de distribución empírica para los tamaños de los eventos de lluvia en Bullileo con MTE entre 1 y 6 horas	102

ÍNDICE DE TABLAS

Tabla 3.1.	Estaciones consideradas en la provincia de Málaga	21
Tabla 3.2.	Estaciones consideradas en la región del Maule (Chile)	25
Tabla 3.3.	Estaciones consideradas en la región del Bío Bío (Chile)	27
Tabla 3.4.	Efemérides para la estación de Málaga Aeropuerto (AEMET, 2012)	35
Tabla 3.5.	Nº de valores que superan las efemérides de la AEMET de cada estación.	35
Tabla 3.6.	Nº de datos iniciales, depurados y descartados.....	39
Tabla 3.7.	Datos descartados en estaciones de la región del Maule (Chile)	42
Tabla 4.1.	Valores críticos de la discordancia (Hosking y Wallis, 1997).....	50
Tabla 4.2.	Valores de los L-momentos de las estaciones de la provincia de Málaga.....	57
Tabla 4.3.	Comportamiento de escala para los datos de las distintas estaciones	62
Tabla 4.4.	Valores de qmax para las distintas estaciones meteorológicas	66
Tabla 4.5.	Identificación de regiones	67
Tabla 4.6.	Regiones formadas mediante Análisis Clúster	69
Tabla 4.7.	Resultados obtenidos de las funciones de distribución.....	70
Tabla 4.8.	Curvas regionales de frecuencia para cada región	71
Tabla 4.9.	Valores de los L-momentos para las estaciones analizadas en la región de Biobío.	72
Tabla 4.10.	Parámetros multifractales obtenidos.	77
Tabla 4.11.	Región Homogénea 1.....	80
Tabla 4.12.	Región Homogénea 2.....	81
Tabla 4.13.	Región Homogénea 3.....	81

Tabla 4.14.	Región Homogénea 4.	82
Tabla 4.15.	Región Homogénea 5.	82
Tabla 4.16.	Valores de las dimensiones fractales D1 y D2, así como del grado de multifractalidad (MD) para las series de datos analizadas (ID) en la región del Maule.....	86
Tabla 4.17.	L-Momentos para las 53 estaciones analizadas.....	89
Tabla 4.18.	Resultados de los tests de homogeneidad para las distintas regiones formadas.	91
Tabla 4.19.	Características de los eventos de lluvia para diferentes TME en Málaga y Bullileo.	96

RESUMEN

Esta tesis doctoral se centra en el análisis multifractal de series históricas de datos de precipitación y en el posterior uso de los parámetros obtenidos para diversas aplicaciones hidrológicas.

Para ello se ha trabajado con datos de lluvia registrados en diversas localidades de España y de Chile, a los que se les han aplicado una serie de procedimientos de control de calidad que garanticen los resultados de análisis posteriores.

La tesis está dividida en dos partes fundamentales. La primera de ella se centra en la formación de regiones homogéneas de precipitación en base a la caracterización multifractal de los datos validados. Para ello se utilizan por un lado, los parámetros multifractales obtenidos tras la aplicación del formalismo multifractal de turbulencia a datos de precipitación diaria en estaciones de la provincia de Málaga (España) y de la región del Biobío (Chile). Por otro lado, se utilizan los valores de la dimensión fractal generalizada obtenidos tras la aplicación del formalismo multifractal del Strange Attractor a las series de lluvia diaria disponibles para la región del Maule (Chile).

En la segunda parte propone una metodología para determinar el tiempo mínimo entre eventos necesario para delimitar eventos de lluvia. Esta metodología se basa en la combinación de los resultados obtenidos tras la aplicación el análisis multifractal a series históricas de datos horarios de precipitación en Málaga (España) y Bullileo (Chile), junto con el comportamiento críticamente autoorganizado de las mismas.

Para todo ello se describen los fundamentos teóricos de las diversas teorías aplicadas, las características más relevantes de los datos de precipitación usados, los resultados obtenidos en cada una de las aplicaciones llevadas a cabo y las conclusiones más relevantes a las que se ha llegado en esta tesis.

ABSTRACT

This PhD thesis is focused in the multifractal analysis of rainfall data series for its use in hydrological purposes. Rainfall data sets from some locations of Spain and Chile have been used after applying some quality control methods in order to guarantee the final results.

The document can be divided into two main parts. The first one shows how to form homogeneous rainfall regions by using the multifractal characterization of the validated rainfall data sets. Firstly, the turbulence formalism was applied to the daily rainfall data series from Malaga region (Spain) and Bibobío region (Chile), and the multifractal parameters were obtained. Afterwards, the generalized fractal dimension values were obtained for the Maule region (Chile) through the Strange Attractor multifractal method.

The second part describes a methodology to determine the minimum inter-event time needed to separate rainfall events. The method is based on the combination of the multifractal results previously obtained for the hourly rainfall data sets in two locations (Málaga, in Spain, and Bullileo, in Chile) and the self-organized criticality behaviour of both data sets.

The theoretical fundamentals of all the theories applied are described in the document, along with the main characteristics of rainfall data used, all the results obtained for each application and the most relevant conclusions of this PhD thesis.

1. INTRODUCCIÓN

La precipitación es una variable compleja que se utiliza en multitud de modelos hidrológicos de cálculo de escorrentía, infiltración, erosión, estimación de curvas intensidad-duración-frecuencia, etc. Actualmente, el conocimiento del comportamiento de la lluvia tanto en el tiempo como en el espacio se orienta fundamentalmente al estudio del cambio climático, a la detección de tendencias, al análisis del aumento de la frecuencia de ocurrencia de eventos extremos de lluvia (Haylock y Nicholls, 2000) o a una mayor presencia de períodos secos (Seager *et al.*, 2007), entre otros.

Para muchos de los estudios citados anteriormente se hace necesario contar con series de datos de una longitud adecuada a los resultados y a las estimaciones que se pretendan. Lamentablemente, la escasez de series de datos de cierta longitud es un problema frecuente en cualquier trabajo de hidrología. El análisis regional de frecuencias se presenta como una alternativa relativamente reciente que permite solucionar la escasez de datos en el tiempo con el aumento de datos en el espacio (Hosking y Wallis, 1997). Esta técnica se basa en el uso de los momentos lineales, introducidos por Hosking (1990, 1992) y utilizados en todos los pasos del método de análisis propuesto por Hosking y Wallis en 1997.

La técnica del análisis regional de frecuencias, cuando se utiliza para la estimación de cuantiles, proporciona valores más precisos que los que podrían obtenerse mediante un análisis local tradicional (Hosking *et al.*, 1985; Lettenmaier y Potter, 1985; Hosking y Wallis, 1988, 1997), siempre y cuando haya sido posible formar regiones totalmente homogéneas respecto a la variable que se esté analizando. Y así aparece uno de los problemas fundamentales a la hora de trabajar con el análisis regional de frecuencias: la agrupación de series de datos (en adelante, estaciones) en grupos o regiones homogéneas.

Existen diversas metodologías para agrupar estaciones en regiones potencialmente homogéneas (e.g. Bonelland y Sumner, 1992; Jingyi y Hall, 2004; Yürekli y Modarres, 2007; Srinivas *et al.*, 2008; García-Marín *et al.*, 2011; Satyanarayana y Srinivas, 2011) y en todas ellas es fundamental definir las características de los diversos individuos que sirvan de base para la agrupación. Al trabajar con datos de lluvia, las características más comunes incluyen parámetros derivados del ajuste a funciones de distribución de

probabilidad (e.g. Easterling, 1989) o bien atributos relacionados con la localización de las propias estaciones meteorológicas (e.g. Guttman, 1993; Moreno y Roldán, 1999). En estudios más recientes (e.g. Satyanarayana y Srinivas, 2011) se ha trabajado también con diversas variables atmosféricas, lo que pone de manifiesto que pueden utilizarse otras características para agrupar series de datos de precipitación procedentes de diversas localizaciones.

Por ello, se propone en esta Tesis, como primer objetivo, el uso de características relacionadas con la invarianza de escala presente en el proceso de la lluvia. El análisis de la invarianza de escala puede abordarse mediante el análisis multifractal del proceso de estudio. En el caso de la lluvia el análisis multifractal se ha estudiado en las última décadas con diferentes propósitos (e.g. de Lima y Grasman, 1999; Castro *et al.*, 2004; Langousis y Veneziano, 2007; Royer *et al.*, 2008; de Lima y de Lima, 2009; Langousis *et al.*, 2009; Valencia *et al.*, 2010; Guadagnini *et al.*, 2012; García-Marín *et al.*, 2013a; Rodríguez *et al.*, 2013; Gires *et al.*, 2014).

Entre las principales ventajas del análisis multifractal se encuentran por un lado la independencia que existe entre el valor de los diferentes parámetros y los datos disponibles, y por otro, el no tener que asumir una función de distribución de probabilidad a la que se ajuste el conjunto de datos analizado.

Existen diversas metodologías para analizar el carácter multifractal de la precipitación, entre las que se encuentra el formalismo Multifractal de Turbulencia (Schertzer y Lovejoy, 1987) y el formalismo del Strange Attractor (e.g. Hentschel y Procaccia 1983; Grassberger 1983; Halsey *et al.* 1986). El primero de ellos analiza el comportamiento multifractal de una serie de datos a partir del método del escalado de momentos (Frisch y Parisi, 1985; Schertzer y Lovejoy, 1987; Over y Gupta, 1994), mientras que el segundo de los formalismos analiza el carácter multifractal a partir de los valores de la dimensión fractal de un conjunto de datos. Cuando diferentes conjuntos de un objeto presentan diferentes dimensiones fractales, el objeto exhibe multifractalidad. En el caso de la lluvia, el ‘objeto’ es un proceso aleatorio. En el caso de la precipitación, la multifractalidad implica que el proceso de la lluvia se comporta estadísticamente igual a pequeña y a gran escala, salvo pequeñas transformaciones (Veneziano y Iacobellis, 2002).

La invarianza de escala no es solo una característica de los procesos multifractales, sino que también está presente en los sistemas críticamente autoorganizados (en adelante sistemas SOC, del inglés self-organized criticality systems). Los sistemas SOC se caracterizan por exhibir leyes potenciales en la distribución del tamaño de sus eventos, lo que implica la ausencia de una escala característica para dichos tamaños. Estos sistemas alcanzan un estado de estabilidad estadística a la que se llega gracias a fluctuaciones temporales caracterizadas por su liberación de energía.

Existen numerosos trabajos que ponen de manifiesto que la teoría SOC puede utilizarse para describir la complejidad de la lluvia (e.g. Andrade *et al.*, 1998; Cowpertwait *et al.*, 2002; Peters y Christensen, 2002; Peters *et al.*, 2002; Telesca *et al.*, 2004; Arakawa, 2006; Bove *et al.*, 2006; Peters y Christensen, 2006; Peters y Neelin, 2006; García-Marín *et al.*, 2007; Peters *et al.*, 2010; Deluca y Corral, 2014) a partir del análisis de la distribución del tamaño de sus eventos (Corral, 2015). Dichos eventos se definen como una secuencia de registros continuos de lluvia distintos de cero y delimitados por períodos sin lluvia (Peters *et al.*, 2002).

En base al análisis de invarianza de escala a partir del carácter multifractal y críticamente autoorganizado de la precipitación, se propone como segundo objetivo de esta Tesis la determinación del tiempo mínimo entre eventos más adecuado para poder separar eventos de lluvia en un determinado lugar.

2. FUNDAMENTOS TEÓRICOS

2.1. La invarianza de escala: fractales y multifractales

La invarianza de escala de un sistema, característica principal de la teoría de los fractales, puede estudiarse mediante el uso de leyes potenciales, caracterizadas por sus exponentes, e indica que determinados rasgos de un sistema son independientes de la escala con la que se analizan. Por tanto, esta teoría es aplicable a sistemas que carecen de una escala característica (García-Marín, 2007).

Un fractal es un objeto o sistema cuya estructura básica, fragmentada o irregular, se repite a diferentes escalas. El término fue propuesto por el matemático Benoit Mandelbrot en 1975 y deriva del latín *fractus*, que significa quebrado o fracturado. Entre sus características se encuentran la auto semejanza, es decir, que cualquier parte arbitraria contiene una réplica exacta del objeto total, y la iteración, que implica que el proceso de construcción del fractal es iterativo (Florez, 2008).

Generalmente, la estructura y organización de un conjunto fractal no hace posible especificar donde se sitúa cada punto que lo compone, de ahí que deba definirse una relación entre las distintas estructuras observadas para varios niveles de resolución, formulada mediante el concepto de dimensión fractal, que describe el comportamiento de escala de las estructuras fractales. Teniendo esto en cuenta, un proceso fractal es aquel en el que el mismo proceso elemental tiene lugar a distintas escalas.

Independientemente de que se hable de objetos, sistemas o procesos, pueden distinguirse los siguientes tipos de fractales:

- Fractales determinísticos: se construyen mediante reglas deterministas y se clasifican en uniescalares y multiescalares, teniendo en cuenta que los uniescalares se obtienen dividiendo un objeto simple en N piezas idénticas, siendo cada una de ellas la copia del objeto original, reducido por el mismo factor $r < 1$. Repitiendo el proceso indefinidamente se obtiene el fractal. Los fractales multiescalares se obtienen también dividiendo en N partes, esta vez no idénticas, tratándose de reducciones del objeto original por diversos factores $r_j < 1$, variando j desde 1 hasta N .

- Fractales aleatorios: generados por leyes no deterministas, presentan aleatoriedad en cada fase de su construcción, de ahí que no tengan auto-similitud entre las partes que lo forman. Esta apariencia es la que los hace útiles para describir determinados fenómenos naturales como las líneas de costas o las nubes.

En la naturaleza existen determinados procesos y sistemas que no pueden estudiarse como conjuntos simples, ya que se caracterizan por distintos niveles de intensidad. Por tanto, debe recurrirse a la teoría multifractal, que facilita el estudio de procesos no lineales con distintos niveles de intensidad. El término “multifractal” fue introducido por U. Frish y G. Parisi en 1985 para hacer referencia a medidas caracterizadas por un conjunto infinito de índices de escala o dimensiones fractales.

El análisis multifractal se usa para caracterizar sistemas o procesos, asignándoles una función llamada espectro multifractal, caracterizada mediante una gama de dimensiones fractales diferentes asociadas a una jerarquía de subconjuntos, cada uno de ellos de carácter fractal. El comportamiento de un sistema multifractal es análogo al comportamiento de tipo cascada observado en el fenómeno de la turbulencia. Las características fundamentales de un fenómeno tipo cascada son la invarianza de escala y la conservación de flujos, desde las escalas mayores a las menores.

En el proceso de desarrollo de una cascada, la estructura de partida se divide en subestructuras, cada vez más pequeñas, de longitud δ , siendo un factor aleatorio el que determina la porción de energía que se transfiere a cada una de las subpartes. Después de n iteraciones del proceso en un espacio de dimensión D , habrá un número λ_D de subestructuras con una longitud y una densidad de flujo.

Los distintos tipos de cascadas y su comportamiento dan lugar a la siguiente clasificación de multifractales (De Lima, 1998):

- Desnudos y vestidos (del inglés ‘bare and dressed’): los multifractales desnudos se obtienen tras un número infinito de pasos, en los que la cascada se desarrolla desde los valores mayores a los menores. En cambio, los multifractales vestidos se obtienen a partir de un proceso físico que considera los valores medios.
- Fuertes y débiles (del inglés ‘hard and soft’): en los multifractales fuertes o duros los momentos estadísticos de mayor orden divergen debido a la presencia

de picos o singularidades. En cambio, en los multifractales débiles o blandos, las fluctuaciones de esos picos son pequeñas, por tanto no ocurre la mencionada divergencia de momentos.

- Microcanónicos y canónicos (del inglés ‘microcanonical and canonical’): una cascada microcanónica es aquella en la que el flujo de energía se conserva exactamente en cada paso, mientras que una cascada canónica tiene lugar cuando se produce sólo una conservación media de la energía.
- Calmados y salvajes (del inglés ‘calm and wild’): la singularidad calmada es aquella que no altera la conservación de la energía en un proceso microcanónico, mientras que la singularidad salvaje es la que sí lo hace.

La caracterización cuantitativa de los fractales se hace mediante el índice de escala o dimensión fractal, donde la escala de medida δ es un parámetro muy importante. Supongamos un objeto de tamaño L definido en un espacio unidimensional. El cociente o razón de escala vendría definido por:

$$\lambda = \frac{L}{\delta} \quad (2.1)$$

Supongamos ahora un conjunto A definido en un espacio D -dimensional. N_λ será el número de cubos que no solapan el lado λ^{-1} , necesarios para cubrir la parte del espacio D -dimensional que contiene al conjunto de interés, cumpliéndose la relación:

$$N_\lambda \approx \lambda^D \quad (2.2)$$

lo que implica que N_λ es proporcional a λ^D cuando $\lambda \rightarrow \infty$. Si $N_{\lambda,A}$ es el número de cubos de lado λ^{-1} sin solapamiento necesarios para completar el conjunto A , se satisface la relación:

$$N_{\lambda,A} \approx \lambda^{D,A} \quad (2.3)$$

donde D_A es la dimensión fractal del conjunto. Por tanto, el conjunto A será fractal si D_A no es un número entero.

La codimensión fractal C vendrá dada entonces por:

$$C = D - D_A \quad (2.4)$$

siendo D la dimensión del espacio en el que se incluye el conjunto A .

2.2. Caracterización de procesos multifractales

La caracterización multifractal es una herramienta eficiente a la hora de describir la estructura temporal de diversos fenómenos (Jiménez-Hornero *et al.*, 2011). Para identificar la multifractalidad de un proceso, existen diversas metodologías, entre las que destacan las que se describen en los siguientes apartados.

2.2.1. El formalismo multifractal de turbulencia: el escalado de momentos

El formalismo multifractal de turbulencia, inicialmente desarrollado por Schertzer y Lovejoy (1987), es una de las metodologías más utilizadas en hidrología (e.g. Schertzer y Lovejoy, 1988; De Lima y Grassman, 1999; De Lima y de Lima, 2009; García-Marín *et al.*, 2015). Este método asume que la variabilidad del proceso en estudio puede modelarse directamente como un proceso estocástico (o aleatorio) de cascada turbulenta (Schertzer y Lovejoy, 1987; Gupta y Waymire, 1993; Over y Gupta, 1994; Lovejoy y Schertzer, 1995). Según este método, la estructura multifractal de un proceso ha de analizarse estudiando el escalado (múltiple) de sus momentos estadísticos.

El escalado de los momentos se analiza a través de la función exponente escaladora de momentos $K(q)$, que ha de satisfacer:

$$\langle \varepsilon_\lambda^q \rangle \approx \lambda^{K(q)} \quad (2.5)$$

donde $\langle \varepsilon_\lambda^q \rangle$ es el momento medio q^{th} de la intensidad del proceso en estudio para una escala λ (que es el cociente entre la longitud total del set de datos analizado y un intervalo temporal cualquiera). Representando en un diagrama doblemente logarítmico los valores de $\langle \varepsilon_\lambda^q \rangle$ como función de λ , podrá distinguirse una zona de apariencia aproximadamente lineal cuya pendiente es una estimación de $K(q)$. Repitiendo este procedimiento para diferentes valores de q , se puede estimar la función $K(q)$ completa.

La función exponente escaladora de momentos $K(q)$ se utiliza para caracterizar el comportamiento de escala (e.g. Svensson *et al.*, 1996). Para procesos (mono) fractales o de escala simple (e.g. Mandelbrot, 1972) la función $K(q)$ frente a q exhibe un comportamiento lineal sin pasar por el origen, mostrando solo un nivel de intensidad para el proceso que se esté analizando. Cuando la función $K(q)$ es lineal y pasa por el origen, el proceso estudiado se describe como auto-similar. Cuando la función exponente escaladora de momentos es no lineal, el proceso es multifractal (Veneziano *et al.*, 2006).

A partir de la función $K(q)$ pueden obtenerse una serie de parámetros multifractales, como el valor de la singularidad máxima γ_{max} y el momento crítico q_{crit} . La función $K(q)$ exhibe una zona lineal una vez superado el valor crítico q_{crit} . La pendiente de esta zona lineal es el orden de singularidad máxima (γ_{max}) presente en el conjunto de datos analizado (e.g. Schertzer y Lovejoy, 1987; Tessier *et al.*, 1993; Lovejoy y Schertzer, 1995; de Lima y de Lima, 2009) y puede determinarse a partir de,

$$\gamma_{max} = \max(K'(q)) \quad (2.6)$$

A partir de la función $K(q)$ y para q igual a 0, puede obtenerse también información de los valores bajos del proceso analizado. La dimensión fractal puede obtenerse también a partir de: $D = 1 + K(0)$. Los valores de $K(0)$ están relacionados con los ‘ceros’ del proceso, de forma que si se obtiene un valor de dimensión fractal cercano a 1 se estaría analizando un set de datos con muy pocos valores nulos (e.g. De Lima y de Lima, 2009; García-Marín *et al.*, 2013).

La caracterización multifractal de un proceso puede también llevarse a cabo de manera teórica a partir de lo que se conoce como la función exponente teórica escaladora de momentos $K(q)_t$. Para obtenerla se aplica el Modelo Multifractal Universal (MMU) que se basa en el formalismo multifractal de turbulencia desarrollado por Schertzer y Lovejoy (1987). La función teórica $K(q)_t$ para momentos $q \geq 0$ y $\alpha \in [0, 2]$ se obtendría según:

$$K(q) - qH = \begin{cases} \frac{C_1}{\alpha - 1} (q^\alpha - q) & \text{para } \alpha \neq 1 \\ C_1 q \ln(q) & \text{para } \alpha = 1 \end{cases} \quad (2.7)$$

Donde α es el índice de Lévy e indica el grado de multifractalidad del proceso (o su desviación respecto a la mono-fractalidad). Su valor oscila entre 0 y 2. El parámetro $C_I \in [0, d]$, donde d es la dimensión del geométrica del proceso analizado, describe la dispersión del valor medio del proceso analizado. El parámetro H es un indicador de la desviación respecto al estado conservativo ($\langle \varepsilon_\lambda \rangle = \lambda^{-H}$). Para procesos conservativos $H = 0$.

La estimación de los parámetros C_I y α puede hacerse aplicando la técnica del Momento de Trazas Doble (MTD) (del inglés Double Trace Moment) (Lavallée *et al.*, 1993). Según este método, la intensidad ε_λ asociada a la menor resolución Λ , ha de elevarse primero a la potencia η , variando la resolución hasta el índice de escala λ y calculando la q -ésima potencia, haciendo el promedio para todos los datos disponibles. La función exponente escaladora de momentos de traza doble $K(q, \eta)$, se obtiene de:

$$\left\langle \left(\varepsilon_\lambda^\eta \right)_\lambda^q \right\rangle = \lambda^{K(q, \eta)} \quad (2.8)$$

Así, $K(q, \eta)$ puede determinarse para valores específicos de q y de η , como la pendiente de la parte lineal obtenida al representar en un gráfico doblemente logarítmico $\left\langle \left(\varepsilon_\lambda^\eta \right)_\lambda^q \right\rangle$ frente a λ . La función $K(q, 1)$, obtenida para $\eta = 1$, y la función exponente escaladora de momentos $K(q)$, son la misma. $K(q, \eta)$ se relaciona con $K(q, 1) \equiv K(q)$ mediante la expresión $K(q, \eta) = K(q\eta, 1) - qK(\eta, 1)$. Mediante esta última expresión y considerando $H = 0$, se obtiene de la ecuación (2.7) que la función $K(q, \eta)$ tiene una dependencia simple de η ,

$$K(q, \eta) = \eta^\alpha K(q) \quad (2.9)$$

De esta forma, el parámetro α se puede estimar como la pendiente de la regresión lineal ajustada al representar el $\log|K(q, \eta)|$ frente a $\log(\eta)$ para valores concretos de q (e.g. Tessier *et al.*, 1996). La estimación de C_I se realiza considerando que $K(q, 1)$ es la interceptación de la regresión lineal con $\log(\eta) = 0$. Así, según la ecuación (2.7) y para $\alpha \neq 1$,

$$C_1 = (\alpha - 1)K(q, 1)/(q^\alpha - q) \quad (2.10)$$

Alternativamente, el parámetro C_1 puede calcularse según la expresión $C_1 = [dK(q)/dq]_{q=1}$.

A partir del valor del exponente β que caracteriza el espectro de energía del proceso conservativo $E(\omega) \approx \omega^{-\beta}$ (siendo ω la frecuencia), el parámetro H se obtendrá de:

$$H = 0.5(\beta - 1 + K(2)) = 0.5(\beta - 1 + C_1(2^\alpha - 2)/(\alpha - 1)) \quad (2.11)$$

2.2.2. El formalismo multifractal del “Strange Attractor”

Los formalismos multifractales están basados en la teoría de las medidas, de forma que las medidas multifractales están relacionadas con la distribución de un determinado valor de un proceso en un determinado soporte geométrico (De Bartolo *et al.*, 2000). El formalismo del Strange Attractor (Hentschel y Procaccia, 1983; Grassberger, 1983; Halsey *et al.*, 1986) trabaja con las dimensiones fractales de los conjuntos geométricos asociados a las singularidades de las medidas.

La dimensión fractal de un conjunto se define como el exponente de escala D_0 ,

$$N(r) = \frac{A}{r^{D_0}} \quad (r \rightarrow \infty) \quad (2.12)$$

Donde $N(r)$ es el número de cajas (del inglés ‘box’) de longitud o lado r , necesarias para cubrir un determinado conjunto, y A es una constante (Mandelbrot, 1982; Feder, 1988). Supongamos que un conjunto se representa por un gran número de puntos. Si todos los puntos se distribuyen de manera uniforme a lo largo del conjunto, la dimensión fractal caracteriza completamente la dimensión del conjunto. Si por el contrario los puntos no se distribuyen de manera uniforme, es posible que la distribución de masas de los puntos varíe. De esta forma, para una determinada longitud de caja r es posible identificar regiones con las mismas medias μ (Feeny, 2000). Para una caja de lado r , la media puede estimarse como $\mu_i = n_i / n$ donde n_i es el número de puntos en la caja, y n el total de puntos. De esta forma puede obtenerse,

$$M_d(q, r) = \sum_{i=1}^N \mu_i^q r^d \quad (2.13)$$

Donde N es el número de cajas que cubren el conjunto analizados; $d = \tau_q$ se conoce como el exponente de masa. Definiendo $Z(q, r) = \sum_{i=1}^N \mu_i^q$ como la función de partición (i.e. Feder, 1988), entonces $Z(q, r) \sim r^{-\tau_q}$ y por lo tanto,

$$\tau_q = \lim_{r \rightarrow 0} \frac{\log Z(q, r)}{\log r} \quad (2.14)$$

τ_q , se puede obtener como la pendiente del segmento lineal de la representación doblemente logarítmica de $Z(q, r)$ frente a r . Para $q \gg 1$, el valor de $Z(q, r)$ viene determinado principalmente por los valores altos de los datos, mientras que la influencia de los datos de valores bajos contribuye en mayor medida a la función de partición para $q \ll -1$ (Kravchenko *et al.*, 1999).

La dimensión fractal generalizada, D_q , para un momento de orden q se define como,

$$D_q = \lim_{r \rightarrow 0} \frac{\log Z(q, r)}{(q-1) \log r} \quad (2.15)$$

En el límite $q \rightarrow 1$, la ecuación (2.15) se reduce a,

$$D_1 = \lim_{r \rightarrow 0} \frac{\sum_{i=1}^N \mu_i \log \mu_i}{\log r} \quad (2.16)$$

Dentro de los valores de la dimensión fractal generalizada, los más usados para describir un determinado proceso son D_0 , D_1 y D_2 . D_0 es la dimensión fractal del conjunto para el que se está realizando el análisis. D_1 es la información de dimensión que describe el grado de heterogeneidad en la distribución de la medida. Según Davis *et al.* (1994), D_1 caracteriza la distribución y la intensidad de las singularidades respecto a la media. Si D_1 disminuye, la distribución de las singularidades estará más dispersa. Por el contrario, si D_1 aumenta las singularidades tendrán valores más bajos exhibiendo una distribución

más uniforme. D_2 es la dimensión fractal de correlación, que está asociada a la función de correlación y determina la distribución media de la medida (Grassberger, 1983; Grassberger y Procaccia, 1983). D_q es una función decreciente respecto a q para una medida distribuida de la multifractalidad (e.g., Saa *et al.*, 2007) donde $D_0 > D_1 > D_2$.

La relación entre el espectro generalizado de dimensiones fractales (también conocido como espectro de Rényi), D_q , y el espectro multifractal, $f(\alpha)$, siendo α el exponente de Lipschitz–Hölder (que cuantifica la fuerza de las singularidades medidas), viene dada por la secuencia de exponentes de masa τ_q (Hentschel y Procaccia, 1983) según la expresión,

$$\tau_q = (q-1)D_q \quad (2.17)$$

El espectro multifractal puede obtenerse a partir de la ecuación (2.15) mediante una transformada de Legendre (ecuación 2.18). Para procesos multifractales, el espectro multifractal presenta forma de parábola invertida. Para medidas monofractales, el valor de α es idéntico para todas las zonas del mismo tamaño y $f(\alpha)$ consiste simplemente en un punto (Kravchenko *et al.*, 1999). El mayor valor del espectro multifractal, $f(\alpha_0)$, se corresponde con el valor de la dimensión fractal D_0 del proceso en estudio. La Figura 2.1. recoge la información más relevante que puede obtenerse a partir del espectro multifractal.

$$\begin{aligned} \alpha_q &= -\frac{d\tau_q}{dq} \\ f(\alpha_q) &= q\alpha_q + \tau_q \end{aligned} \quad (2.18)$$

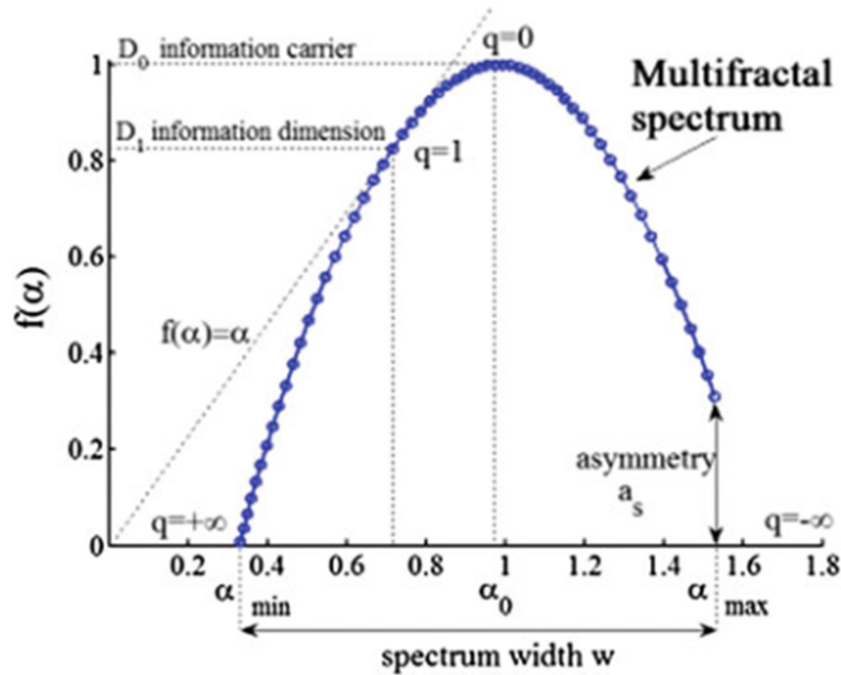


Figura 2.1. Representación esquemática de los diferentes parámetros del espectro multifractal (Figura obtenida de Krzyszczak J., *et al.* 2017)

2.3. La teoría de la criticalidad autoorganizada

La criticalidad autoorganizada (del inglés *self-organized criticality*), en adelante SOC, un concepto introducido por Bak *et al.* (1987). Sirve para explicar algunos patrones complejos, como estructuras fractales y eventos catastróficos.

En ocasiones en la naturaleza, el comportamiento complejo de un sistema formado por varios componentes, implica su evolución hacia un estado ‘crítico’, donde pequeñas alteraciones pueden desencadenar una serie de eventos de cualquier tamaño, conocidos como avalanchas. La mayoría de los cambios se producen mediante eventos catastróficos y sin la intervención de ningún agente externo. El estado del sistema se establece simplemente por las interacciones entre los elementos individuales que lo componen: el estado crítico es autoorganizado.

Los sistemas que siguen un comportamiento críticamente autoorganizado pueden describirse mediante leyes potenciales que arrojan una gran información del proceso

analizado (Corral y Deluca, 2013). Los sistemas críticamente autoorganizados siguen leyes potenciales, pero sin embargo, la existencia de una ley potencial para describir un proceso no garantiza su criticalidad autoorganizada (Deluca y Corral, 2014). Es importante también resaltar el hecho de que si el comportamiento de ley potencial no existe para un determinado rango de valores, esto no quiere decir que no exista comportamiento de escala o criticalidad autoorganizada (e.g. Deluca y Corral, 2014).

Una ley potencial (Bak, 1997) implica que una determinada cantidad N puede expresarse como potencia de otra cantidad x ,

$$n(x) \propto x^{-\tau} \quad (2.19)$$

Donde τ es la pendiente en un gráfico doblemente logarítmico. Las distribuciones potenciales presentan propiedades estadísticas singulares, como la invarianza de escala y la divergencia de momentos (Corral y Deluca, 2013). Si un conjunto de datos sigue una distribución potencial, no presenta una escala típica y prevalece la invarianza de escala.

Para analizar $n(x)$ puede utilizarse una función de distribución empírica de probabilidad (e.g. Lovejoy y Schertzer, 1995; Fraedrich y Larnder, 1993; Tessier *et al.*, 1996; Sivakumar, 2001; Peters *et al.*, 2010). Si la cola de la distribución de probabilidad de x sigue una ley potencial

$$\Pr(X > x) \propto x^{-\tau} \quad (2.20)$$

Para valores altos de x , la serie de valor se caracteriza por su intermitencia hiperbólica (Svensson *et al.*, 1996). Una forma útil de analizar $n(x)$ es mediante el uso de la función de distribución acumulada $\bar{N}(x)$ (Andrade *et al.*, 1998) dividida entre x ,

$$N(x) = \frac{\bar{N}(x)}{x} \propto \frac{1}{\tau - 1} x^{-\tau} \left[1 - \left(\frac{x}{Max} \right)^{\tau - 1} \right] \quad (2.21)$$

donde Max es el evento máximo presente en el conjunto de datos analizado.

Si se tiene en cuenta que la invarianza de escala y la divergencia de momentos está presente tanto en los sistemas multifractales como en los críticamente autoorganizados, puede establecerse relación entre la función $K(q)$ (ecuación 2.5) y la distribución

hiperbólica de los eventos extremos (ecuación 2.21), siendo el exponente τ de la ecuación (2.21) una estimación del momento crítico q_{crit} .

Según autores como Pandey *et al.* (1998) y teniendo en cuenta que la función empírica $K(q)$ sigue un comportamiento lineal para valores de q mayores que q_{crit} , puede establecerse que,

$$\langle \varepsilon_\lambda^q \rangle \rightarrow \infty \quad \text{para } q > q_{crit} \quad (2.22)$$

La divergencia de los momentos estadísticos corresponde a la caída hiperbólica de la distribución de probabilidad (ecuación 2.23) en la que s es un límite,

$$Pr(\langle \varepsilon_\lambda^q \rangle \geq s) \approx s^{q_{crit}} \quad \text{para } s \gg 1 \quad (2.23)$$

La divergencia de los momentos estadísticos de mayor orden (ecuación 2.22) y la distribución potencial equivalente de los extremos (ecuaciones 2.20 y 2.23), son consecuencias directas de la multifractalidad.

3. DATOS DE LLUVIA

3.1. Zonas de estudio

Para el desarrollo de esta Tesis se han utilizado datos de precipitación de distinta resolución temporal ubicados en zonas de España (Figura 3.1) y Chile (Figuras 3.2 y 3.3) que, aunque claramente distantes, presentan un clima común marcado por su carácter Mediterráneo. En el caso concreto de España, este estudio se centra en la provincia de Málaga y en el caso de Chile se centra en las regiones VII (Región del Maule), y VIII (Región del Biobío).

La provincia de Málaga está situada en Andalucía, concretamente en el extremo oeste del Mar Mediterráneo y en el Sur de la Península Ibérica. Su capital se encuentra a unos 100 Km al este del Estrecho de Gibraltar, entre las provincias de Granada, al este, y Cádiz, al oeste (Figura 3.1). El patrón general de lluvias de la provincia de Málaga se caracteriza por lluvias frontales, principalmente entre los meses de Octubre y Mayo, y tormentas de tipo convectivo a finales de verano y principios de otoño. Durante los meses de Julio y Agosto las precipitaciones son prácticamente inexistentes. La precipitación media anual de Málaga es de 524 mm y la temperatura media anual es de 18°C.

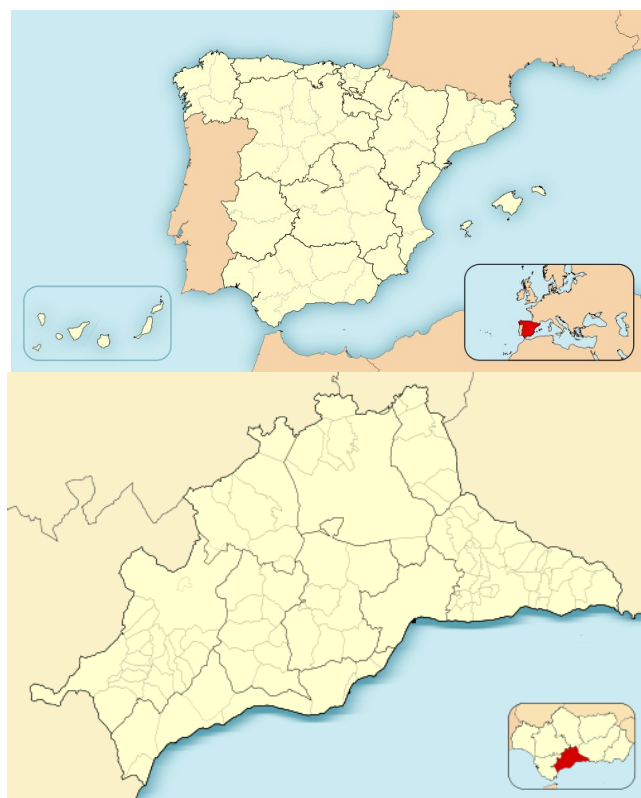


Figura 3.1. Mapa de situación de la provincia de Málaga en España

Por su parte, la Región del Maule (Figura 3.2) se encuentra en la zona semiárida de Chile, con precipitaciones anuales que oscilan entre los 600 mm de los valles y los 2.300 mm en la cordillera Maulina. Los veranos son comúnmente cálidos y secos, al contrario de los inviernos, que suelen ser lluviosos y frescos, con frecuentes heladas influenciadas por el efecto continental. El relieve de esta región es muy variado, lo que condiciona su climatología, pudiéndose diferenciar cuatro zonas geomorfológicas: cordillera de los Andes, Depresión Intermedia, cordillera de la Costa y planicies Litorales.



Figura 3.2. Región del Maule (Chile)

La Región de Biobío (Figura 3.3) es otra de las quince regiones en las que se encuentra dividido Chile. Se sitúa entre los $36^{\circ} 00'$ y $38^{\circ} 30'$ de latitud sur y entre los $71^{\circ} 00'$ de longitud oeste hacia el océano Pacífico, limitando al norte con la Región de Maule, al sur con la Región de la Araucanía, al este con la República Argentina y al oeste con el océano Pacífico.

El clima de esta región es mediterráneo, sin embargo, al sur la influencia mediterránea cambia a oceánica. Se desarrolla en una franja longitudinal, pero en el interior y el sur se deja sentir la influencia mediterránea más lluviosa, dejando más de 2.400 mm de precipitaciones concentradas en invierno.



Figura 3.3. Región del Bio Bio (Chile)

3.1.1. Estaciones seleccionadas

Para los distintos estudios llevados a cabo en esta Tesis, se han seleccionado datos de precipitación de distintas duraciones en las zonas descritas en el apartado anterior.

Así, en el caso de España, los datos de lluvia seleccionados han sido:

- Serie de precipitaciones máximas diarias para setenta y dos estaciones meteorológicas situadas en la provincia de Málaga (Andalucía-España), y suministradas por la Cuenca Mediterránea Andaluza. La Tabla 3.1 recoge la información detallada de su emplazamiento, así como la longitud de las series disponibles. Por su parte, la Figura 3.4 muestra la distribución espacial de estas estaciones en la provincia.

Tabla 3.1. Estaciones consideradas en la provincia de Málaga

ID	Nombre	Periodo	Latitud (N°)	Longitud (W°)	Elevación (m)
1	Agujero	1952-1999	36° 46' 25''	4° 25' 57''	100
2	Alcaucín Cortijo	1947-1999	36° 51' 50''	4° 09' 31''	220
3	Alcaucín Forestal	1946-1999	36° 54' 04''	4° 07' 46''	508
4	Alfarnate	1941-1999	36° 59' 40''	4° 15' 36''	880
5	Alhaurín El Grande	1964-1999	36° 37' 40''	4° 41' 46''	400
6	Aljaima	1946-1999	36° 44' 04''	4° 40' 21''	50
7	Almargen	1945-1999	37° 00' 10''	5° 01' 16''	510
8	Almogía Los Llanes	1949-1999	36° 49' 34''	4° 32' 26''	363
9	Alora	1946-1999	36° 49' 09''	4° 41' 56''	110
10	Alozaina	1944-1999	36° 43' 40''	4° 51' 26''	386
11	Alpandeire	1940-1999	36° 38' 00''	5° 12' 6''	695
12	Antequera El Águila	1943-1999	37° 01' 45''	4° 33' 36''	460
13	Archidona	1936-1999	37° 05' 40''	4° 23' 07''	598
14	Arriate	1947-1999	36° 48' 04''	5° 08' 26''	7
15	Benahavis	1952-1999	36° 31' 25''	5° 02' 46''	190
16	Benalmádena	1964-1999	36° 35' 42''	4° 34' 24''	240
17	Benamargosa	1942-1999	36° 50' 04''	4° 11' 31''	96
18	Benamocarra	1942-1999	36° 47' 25''	4° 09' 36''	179
19	Benaolán CP	1959-1999	36° 41' 15''	5° 16' 51''	920
20	Bobadilla Estación	1944-1999	37° 02' 09''	4° 43' 36''	384
21	Borregos	1949-1999	37° 04' 55''	4° 37' 07''	410
22	Buitreras CE	1941-1999	36° 32' 20''	5° 22' 46''	212
23	Buitreras Presa	1943-1999	36° 34' 45''	5° 20' 06''	334
24	Campillos	1945-1999	37° 02' 50''	4° 51' 46''	460

ID	Nombre	Periodo	Latitud (N°)	Longitud (W°)	Elevación (m)
25	Canillas Aceituno	1942-1999	36° 52' 26''	4° 04' 57''	645
26	Cartajima	1943-1999	36° 38' 47''	5° 09' 15''	846
27	Cartama Estación	1948-1999	36° 44' 04''	4° 36' 21''	35
28	Casabermeja	1945-1999	36° 53' 29''	4° 25' 46''	547
29	Casabermeja V. Pineda	1947-1999	36° 52' 25''	4° 26' 60''	693
30	Casapalma	1959-1999	36° 56' 50''	4° 20' 39''	880
31	Casarabonela Forestal	1946-1999	36° 47' 00''	4° 50' 26''	494
32	Casares	1945-1999	36° 26' 35''	5° 16' 26''	415
33	Chorro Estación	1948-1999	36° 54' 30''	4° 45' 36''	225
34	Coín	1943-1999	36° 39' 40''	4° 45' 31''	209
35	Colmenar	1943-1999	36° 54' 10''	4° 20' 6''	680
36	Cómpeta	1942-1999	36° 50' 00''	3° 58' 31''	636
37	Conde de Guadalhorce	1943-1999	36° 56' 00''	4° 48' 06''	325
38	Contaderas Forestal	1956-1999	36° 50' 40''	4° 23' 16''	630
39	Corchado Central	1942-1999	36° 29' 45''	5° 24' 31''	98
40	Cuevas Becerro	1945-1999	36° 52' 40''	5° 02' 46''	735
41	El Burgo	1943-1999	36° 47' 20''	4° 56' 46''	591
42	Fuente Piedra	1949-1999	37° 08' 05''	4° 43' 47''	440
43	Fuente Piedra Herriza	1964-1999	37° 07' 00''	4° 43' 46''	425
44	Gobantes Vivero	1945-1999	36° 56' 20''	4° 47' 36''	360
45	Humilladero	1946-1999	37° 06' 50''	4° 51' 16''	448
46	Hundidero Pto. Sapo	1954-1999	36° 50' 50''	4° 47' 46''	870
47	Istán	1942-1999	36° 35' 00''	4° 56' 56''	310
48	Jimena Libar CE	1946-1999	36° 39' 40''	5° 16' 51''	429

ID	Nombre	Periodo	Latitud (N°)	Longitud (W°)	Elevación (m)
49	Las Mellizas	1946-1999	36° 52' 00''	4° 43' 16''	140
50	Málaga Oficina	1961-1999	36° 43' 20''	4° 24' 17''	31
51	Marbella Inst. Laboral	1946-1999	36° 30' 29''	4° 53' 06''	20
52	Mijas Faro Cal	1942-1999	36° 30' 29''	4° 37' 60''	20
53	Moclinejo	1940-1999	36° 46' 15''	4° 15' 16''	451
54	Montejaque CE	1942-1999	36° 45' 10''	5° 14' 26''	530
55	Nerja	1947-1999	36° 44' 50''	3° 52' 36''	21
56	Ojen	1948-1999	36° 33' 50''	4° 51' 21''	335
57	Parauta Taramal	1947-1999	36° 42' 35''	5° 01' 36''	1190
58	Parchite	1946-1999	36° 48' 45''	5° 06' 36''	755
59	Peña Enamorados	1943-1999	37° 03' 25''	4° 30' 01''	470
60	Periana	1939-1999	36° 55' 40''	4° 11' 31''	547
61	Pizarra	1946-1999	36° 46' 00''	4° 42' 26''	85
62	Rincón de la Victoria	1964-1999	36° 43' 00''	4° 17' 21''	5
63	Riogordo	1945-1999	36° 54' 54''	4° 17' 36''	400
64	Ronda CE	1940-1999	36° 44' 30''	5° 10' 16''	600
65	SP Alcántara	1946-1999	36° 28' 39''	4° 59' 36''	25
66	Tolox Millanas	1964-1999	36° 41' 25''	4° 51' 31''	200
67	Torrox	1944-1999	36° 45' 29''	3° 57' 06''	145
68	Vegueta Grama	1943-1999	36° 49' 10''	3° 51' 46''	410
69	Vélez Málaga	1940-1999	36° 47' 00''	4° 06' 16''	165
70	Villanueva de Tapias	1950-1999	37° 11' 01''	4° 20' 06''	670
71	Viñuela	1942-1999	36° 51' 05''	4° 08' 06''	130
72	La Yedra	1946-1999	36° 59' 00''	4° 27' 41''	780

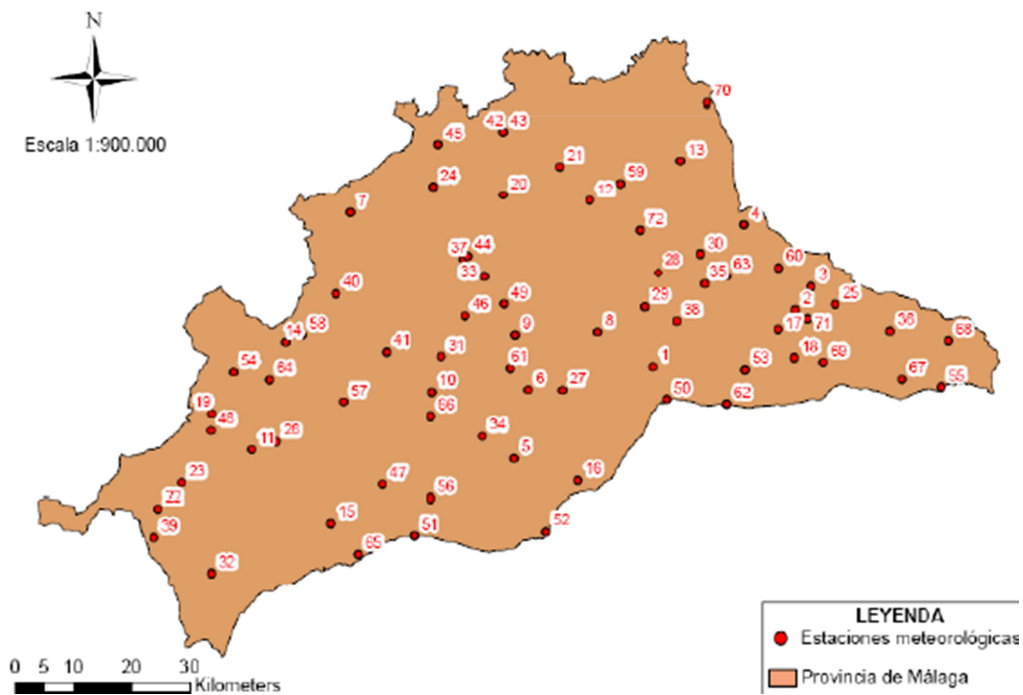


Figura 3.4. Distribución espacial de las estaciones consideradas en la provincia de Málaga.

- Serie de datos horarios de precipitación para una longitud de 30 años (1981-2010, ambos incluidos) de la estación meteorológica “Málaga Aeropuerto”, suministrados por la Agencia Estatal de Meteorología (AEMET).

En el caso de Chile, se ha trabajado con los siguientes datos de lluvia:

- Por un lado, con las series de precipitaciones diarias de cincuenta y tres estaciones meteorológicas situadas en la Región del Maule, suministradas por la Dirección General de Aguas (DGA) del Ministerio de Obras Públicas del Gobierno de Chile. La Tabla 3.2 recoge los detalles de su ubicación mientras que la Figura 3.5 muestra su distribución espacial en la región del Maule.

Tabla 3.2. Estaciones consideradas en la región del Maule (Chile)

ID	Nombre	Periodo	Latitud (S)	Longitud (W)	Elevación (m)
1	Agua Fría	1993-2013	35° 18' 47"	71° 05' 54"	560
2	Ancoa Embalse	1957-2013	35° 54' 38"	71° 17' 45"	421
3	Armerillo	1948-2013	35° 42' 04"	71° 04' 38"	492
4	Bullileo Embalse	1930-2013	36° 17' 06"	71° 24' 51"	600
5	Colbún Maule Sur	1961-2013	35° 37' 27"	71° 24' 08"	280
6	Colorado	1963-2013	35° 38' 17"	71° 15' 38"	420
7	Constitución	1992-2013	35° 19' 27"	72° 24' 32"	10
8	Curicó	1971-2013	34° 58' 52"	71° 14' 10"	195
9	Digua Embalse	1956-2013	36° 15' 21"	71° 32' 53"	390
10	El Álamo	1994-2013	36° 06' 46"	72° 25' 17"	180
11	El Durazno	1992-2013	35° 29' 33"	71° 19' 06"	275
12	El Guindo	1964-2013	35° 15' 28"	71° 19' 26"	250
13	El Manzano	1976-2013	34° 57' 48"	70° 55' 04"	574
14	Fundo El Peral	1966-1986	35° 24' 02"	71° 47' 00"	110
15	Fundo El Radal	1992-2013	35° 25' 08"	71° 02' 35"	685
16	Gualleco	1961-2013	35° 14' 38"	71° 58' 48"	100
17	Hornillo	1962-2013	35° 52' 02"	71° 07' 02"	810
18	Huapi	1969-2013	35° 29' 11"	71° 17' 35"	250
19	Huerta Maule	1992-2013	35° 39' 41"	71° 56' 46"	218
20	Juan Amigo	1992-2013	36° 04' 33"	71° 23' 27"	460
21	La Estrella	1992-2013	35° 46' 57"	72° 11' 13"	200
22	La Sexta	1992-2013	36° 06' 46"	71° 36' 56"	229
23	Liguay	1975-2013	35° 56' 52"	71° 41' 03"	104
24	Linares	1979-2013	35° 50' 17"	71° 35' 43"	157
25	Lontue	1976-2013	35° 02' 32"	71° 17' 26"	199
26	Los Queñes	1931-2013	35° 00' 03"	70° 48' 43"	663
27	Melozal	1951-2013	35° 47' 08"	71° 45' 59"	96

ID	Nombre	Periodo	Latitud (S)	Longitud (W)	Elevación (m)
28	Monte Oscuro	1994-2013	35° 07' 27"	70° 58' 29"	632
29	Nirivilo	1961-2013	35° 32' 20"	72° 05' 29"	200
30	Parral	1964-2013	36° 11' 16"	71° 49' 42"	175
31	Pencahue	1987-2013	35° 22' 21"	71° 49' 57"	55
32	Potrero Grande	1975-2013	35° 11' 00"	71° 05' 52"	445
33	Putú	1992-2013	35° 13' 06"	72° 17' 00"	36
34	Quella	1961-2013	36° 03' 26"	72° 05' 21"	130
35	Río Ancoa Morro	1999-2013	35° 54' 31"	71° 17' 53"	402
36	Río Claro Rauquén	1999-2013	35° 27' 09"	71° 43' 60"	64
37	Río Loncomilla	2001-2013	35° 37' 01"	71° 46' 04"	68
38	Río Longavi	2001-2013	36° 13' 49"	71° 27' 25"	449
39	Río Mataquito	2001-2013	34° 59' 04"	72° 00' 36"	20
40	Río Maule Armerillo	2001-2013	35° 42' 22"	71° 06' 50"	470
41	Río Maule Salto	2003-2013	35° 53'	71° 01' 09"	730
42	Río Maule Forel	2001-2013	35° 24' 25"	72° 12' 30"	30
43	Río Palos	2001-2013	35° 16' 28"	71° 00' 56"	600
44	Río Teno	1999-2013	34° 59' 46"	70° 49' 14"	647
45	San Javier	1970-2013	35° 35' 42"	71° 39' 26"	135
46	San Manuel	1956-2013	36° 21' 29"	71° 38' 58"	270
47	San Rafael	1992-2013	35° 18' 23"	71° 31' 24"	152
48	Santa Susana	1985-2013	34° 54' 41"	71° 02' 07"	410
49	Talca	1964-1982	35° 25' 10"	71° 39' 38"	110
50	Talca UC	1982-2013	35° 26' 09"	71° 37' 11"	130
51	Tutuvén	1978-2013	35° 53' 48"	72° 22' 25"	179
52	Vilches Alto	1992-2013	35° 35' 35"	71° 05' 13"	1058
53	Villa Prat	1992-2013	35° 05' 49"	71° 36' 50"	90

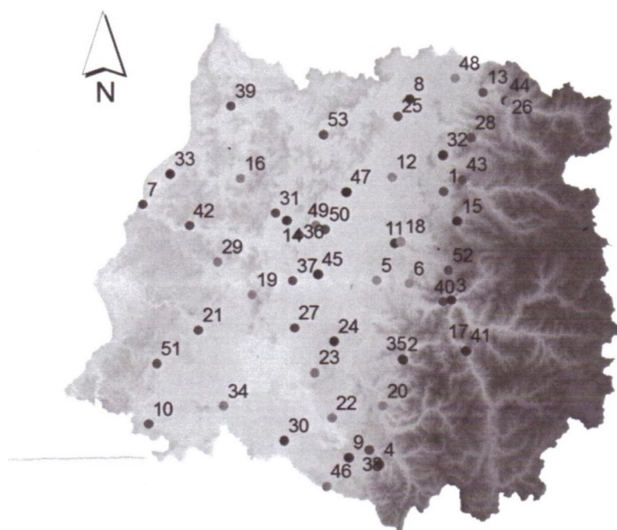


Figura 3.5. Distribución espacial de las estaciones consideradas en la región del Maule (Chile)

Dentro de esta región, se ha trabajado también con la serie de datos horarios de precipitación para una longitud de 39 años (1971-2009, ambos incluidos) de la estación meteorológica “Bullileo”. Estos datos fueron aportados por el Centro Tecnológico de Hidrología Ambiental (CTHA), de la Universidad de Talca. Bullileo se encuentra en la Cordillera de los Andes. En esta zona, los meses más lluviosos son Mayo, Junio y Julio y el periodo más seco se concentra en verano. Las lluvias son, principalmente, de origen frontal, con un promedio anual de 2.105 mm.

- Finalmente, se ha trabajado también con las series de precipitaciones diarias de setenta y cuatro estaciones meteorológicas ubicadas en la Región Biobío de Chile, datos suministrados también por la citada Dirección General de Aguas. El nombre y el identificador de estas estaciones se muestra en la Tabla 3.3.

Tabla 3.3. Estaciones consideradas en la región del Bío Bío (Chile)

ID	Nombre	ID	Nombre
1	Andalien	38	Las Trancas
2	Caman	39	Los Ángeles
3	Canal de la Luz	40	Mangarral

ID	Nombre	ID	Nombre
4	Cancha los Litres	41	Mayulermu
5	Cañete	42	Millauquen
6	Caracol	43	Mulchen
7	Carampangue	44	Nueva Aldea
8	Cerro El Padre	45	Parque Nahuelbuta
9	Cerro Verde	46	Pemuco
10	Chillán Sendos	47	Pilguen
11	Chillán Viejo	48	Portezuelo
12	Chillancito	49	Quilaco
13	Cholguán	50	Quillailleo
14	Coelemu	51	Rafael
15	Coihueco Embalse	52	Rio Biobío Coihue
16	Concepción DGA	53	Rio Biobío Desembocadura
17	Concepción Edif MOP	54	Rio Biobío Esperanza
18	Contulmo	55	Rio Biobío Llauquen
19	Curanilahue	56	Rio Biobío Longitudinal
20	Dichato	57	Rio Biobío Rucalhue
21	Diguillín	58	Río Curanilahue Curanilahue
22	Duqueco	59	Rio Diguillín San Lorenzo
23	Embalse Pangue	60	Río Duqueco Cerrillos
24	Embalse Ralco	61	Río Ítata Coelemu
25	Estero Bellavista Tome	62	Río Laja Salto
26	Estero Hualqui Desembocadura	63	Río Laja Tucapel
27	Estero Nonguen	64	Río Leiva Puentecontulmo
28	Estero Plegarias	65	Río Ñuble San Fabián
29	Estero Quilque	66	San Agustín Puñual
30	Fundo Atacalco	67	San Carlos Puren
31	Fundo Bella Rosa	68	San Fabián
32	General Cruz	69	San José Munilque
33	La Punilla	70	San Lorenzo
34	Laja	71	San Lorenzo Biobío
35	Las Achiras	72	Tirua
36	Las Cruces	73	Trupán
37	Las Pataguas	74	Tucapel

3.2. Control de Calidad

3.2.1. Métodos

En los últimos años se han desarrollado distintos métodos para la validación de datos meteorológicos dentro de los procedimientos enmarcados en los sistemas de control de calidad. Un sistema de control de calidad no sólo se basa en la validación de los registros meteorológicos sino en un conjunto de actividades y decisiones que aseguran la fiabilidad de los registros obtenidos de las estaciones meteorológicas. La finalidad del control de calidad es, por tanto, la detección de posibles errores en el conjunto de datos obtenidos de las estaciones meteorológicas con el fin de hacer un uso óptimo de los mismos (Estévez, 2008).

Existen, básicamente, dos categorías dentro de los procedimientos de control de calidad: los procedimientos que analizan datos de una sola estación y aquellos que requieren datos de varias estaciones.

Para la validación de los datos diarios de precipitación disponibles se ha seguido la metodología de Hubbard *et al.*, (2005), basada en decisiones estadísticas, diseñando un Test de Rango fijo, un Test de Rango dinámico y un Test de Persistencia. Este tipo de tests tienen la gran ventaja de adecuarse a las necesidades de la red de estaciones meteorológicas, además de tener la capacidad de aumentar su eficacia a medida que las series de datos son mayores, ya que utilizan agregados estadísticos que dependen del número de registros (Estévez, 2008). Igualmente, se ha aplicado un test de consistencia espacial y un procedimiento de inspección manual de los datos.

3.2.1.1. Test de Rangos o Intervalos

Con este tipo de test se comprueba si una determinada observación se encuentra dentro de un rango predeterminado que puede ser fijo o dinámico, lo que da lugar a dos tipos de test, el Test de Rango fijo y el Test de Rango dinámico o ‘*upper and lower threshold test*’ (Estévez *et al.*, 2011).

El primer procedimiento de validación diseñado para verificar la calidad de los datos de precipitación es un Test de Rango fijo basado en que la precipitación no puede registrar

ningún dato negativo. Igualmente el límite superior no puede sobrepasar los 508 mm, límite marcado por Shafer *et al.*, (2000),

$$0 \leq P \leq 508mm \quad (3.1)$$

El siguiente procedimiento de validación, basado también en la comprobación del dato con un valor límite específico, utiliza los límites marcados por las efemérides registradas en la estación meteorológica más cercana (Shafer *et al.*, 2000; Feng *et al.*, 2004). El test compara, por tanto, los valores límite con los valores extremos de precipitación. Si el registro de precipitación diaria supera el valor máximo registrado se puede considerar, en principio, como un *outlier* o valor anómalo dentro de la serie de datos, que deberá inspeccionarse posteriormente para determinar si realmente corresponde a un evento extremo para su consideración como nueva efeméride.

3.2.1.2. Tests basados en decisiones estadísticas

Cuando se dispone de gran cantidad de datos iniciales es necesario aplicar una serie de procedimientos de validación algo más complejos, para lo cual se sigue la metodología de Hubbard *et al.*, (2005) basada en decisiones estadísticas.

Este tipo de procedimientos tienen la ventaja de ser unos tests que se adecuan a las necesidades y características de cada red de estaciones meteorológicas y tienen la capacidad de aumentar su eficacia a medida que las series temporales son mayores, ya que utilizan agregados estadísticos que dependen del número de registros.

Básicamente este procedimiento consiste en aplicar un conjunto de algoritmos a cada registro y el proceso de validación acepta el dato, etiquetándolo como verdadero, o lo rechaza, marcándolo como valor anómalo u *outlier*.

Esta hipótesis de verificar cada registro y la decisión estadística de aceptarlo o marcarlo como potencialmente erróneo puede generar dos tipos de errores. Así pues, si el dato es válido y el test lo acepta o si el dato es no válido y el test lo rechaza, se puede afirmar que el test está funcionando correctamente. Cuando un dato válido es rechazado se produce un error Tipo I y si el dato es no válido y el test lo acepta se produce un error Tipo II (Estévez, 2008).

El *upper and lower threshold test* o Test de Rango Dinámico tiene como objetivo la comprobación de que un determinado registro, se sitúa dentro de un rango específico definido para cada mes del año, tal y como se muestra a continuación en la siguiente ecuación (Hubbard *et al.*, 2005),

$$\bar{x} - f\sigma_x \leq x \leq \bar{x} + f\sigma_x \quad (3.2)$$

donde x es el registro diario de precipitación que se va a validar, \bar{x} representa el promedio diario (p.ej. media de los valores diarios de precipitación del mes de Enero, número de años disponibles x 31 días para Enero), f (factor) es un parámetro cuyo valor oscila entre 0,2 y 5 y σ_x es la desviación típica de los valores diarios de x para ese mes en cuestión.

El límite superior de dicho rango viene dado por el valor de la media correspondiente más f veces su desviación típica. Así, el límite inferior viene dado por el valor de la media correspondiente menos f veces su desviación típica.

3.2.1.3. Test de Persistencia

El Test de Persistencia es un procedimiento basado en la consistencia temporal, cuya finalidad es la comprobación de la escasa o nula variabilidad de los registros que se pretenden validar.

Hubbard *et al.* (2005) desarrollaron este test, basado en la hipótesis de que cuando un sensor falla siempre registra un valor constante, de manera que la desviación típica de la serie evaluada será menor, pudiendo ser cero si se completa el periodo de medida. En otros casos, cuando el sensor trabaja intermitentemente, se registrarán valores razonables y otros cercanos a cero, obteniéndose una desviación típica muy elevada. Esto implica que cuando la variabilidad se salga fuera de unos límites determinados, según la Ecuación (3.3), se generará una alerta para todos los datos comprendidos en el periodo de estudio y serán etiquetados como sospechosos.

$$\bar{\sigma}_j - f\sigma_{j,\sigma} \leq \sigma \leq \bar{\sigma}_j + f\sigma_{j,\sigma} \quad (3.3)$$

donde σ es la desviación típica de los valores de precipitación para cada mes (j) y año, $\sigma_{j,\sigma}$ es la desviación típica de σ para cada mes en cuestión y $\bar{\sigma}_j$ es el valor medio de esas desviaciones típicas para cada mes. Igual que en los casos anteriores se ha estudiado la relación existente entre los distintos valores de f y el porcentaje de errores generados.

3.2.1.4. Test de Consistencia Espacial, SRT

El SRT (del inglés Spatial Regression Test) desarrollado por Hubbard *et al.* (2005), comprueba que los datos de la variable en cuestión se encuentran en un intervalo de confianza formado por los datos de una serie de referencia que se crea a partir de una serie de estaciones vecinas. Se selecciona un número de estaciones (M), que serán las más cercanas a la estación candidata. Este número M se fija concretamente en 8 estaciones meteorológicas (Estévez *et al.*, 2011b). A continuación, se selecciona un número de estaciones definitivas, de las marcadas previamente, para lo cual se realiza una regresión lineal de cada estación vecina con la estación candidata, para la serie de datos:

$$P_i = a_i + b_i P_v \quad (3.4)$$

Donde a_i y b_i son los coeficientes de cada regresión lineal, P_v son los registros diarios de cada estación vecina y P_i son los valores medidos de precipitación en la estación candidata. Para cada regresión lineal se calcula el error estándar de la estimación (s_i).

Los valores ponderados de la estimación de la serie de referencia (P') se obtienen utilizando el error estándar de la estimación, también conocido como la raíz cuadrada del error cuadrático medio (RMSE):

$$P' = \left(\sum_{i=1}^N P_i^2 s_i^{-2} \right)^{0,5} \left(\sum_{i=1}^N s_i^{-2} \right)^{-0,5} \quad (3.5)$$

Donde N es el número de estaciones utilizadas en la estimación de la serie de referencia. En este caso, el criterio de selección de las estaciones definitivas es el de aquellas que

tengan un menor RMSE. Siguiendo la metodología citada anteriormente N se va a fijar en 5, teniendo en cuenta siempre que $N < M$.

El error estándar de la estimación ponderado (s') se calcula de la siguiente forma:

$$\frac{1}{s'^2} = N^{-1} \sum_{i=1}^N s_i^{-2} \quad (3.6)$$

A continuación, se establecen los intervalos de confianza para la aplicación de este test a partir de s' , donde se comprueba si el valor registrado en la estación candidata (x) está dentro de estos límites:

$$P' - fs' \leq P_i \leq P' + fs' \quad (3.7)$$

Donde P' es el valor de la serie de referencia, f es un factor que toma distintos valores y P_i es el dato de la variable que se está validando, en este caso, precipitación. Si el registro cumple esta condición, el dato en cuestión pasa satisfactoriamente este test de consistencia espacial. En caso contrario, es decir, si el dato no pasa el test, es rechazado y marcado como un outlier o dato potencialmente erróneo, que tendrá que ser verificado posteriormente. En función del valor que se le asigne a f el porcentaje de datos marcados irá variando.

3.2.1.5. Test de Consistencia espacial SRT Modificado

La aplicación del test de consistencia espacial (SRT) de Hubbard *et al.*, (2005) modificado se basa en el uso de regiones homogéneas previamente definidas. En esta Tesis se utilizarán las regiones homogéneas previamente establecidas en García-Marín *et al.* (2011).

En este caso se dispone de P' , valor de referencia construido en base a las estaciones vecinas, que se va a obtener teniendo en cuenta que N es el número de estaciones que forman la región homogénea.

3.2.1.6. *Inspección Manual*

En el proceso de validación siempre existe una última etapa de inspección visual y análisis manual, que debe ser efectuado por personal debidamente cualificado. Básicamente se trata de visualizar los datos considerados inicialmente como sospechosos, tras la aplicación de los distintos test de validación, para que definitivamente sean admitidos o no como datos erróneos o, en su caso, darlos por válidos y tratarlos como una nueva efeméride meteorológica (Estévez, 2008).

En esta última fase tiene especial importancia el estudio de los eventos extremos (efemérides) que, a veces, pueden darse en nuestra climatología, dando lugar a valores tan destacados que, inicialmente, son considerados como errores potenciales por los procedimientos de validación. En resumen, cabe destacar que la inspección manual resulta esencial en la etapa final de validación de datos meteorológicos.

3.2.2. Resultados

Se describen a continuación los resultados obtenidos al aplicar los diferentes tests a los datos disponibles en las estaciones de España y de Chile.

3.2.2.1. *Resultados de validación en las estaciones de la provincia de Málaga (España)*

La validación de datos de la serie de precipitaciones máximas diarias de las setenta y dos estaciones meteorológicas de la provincia de Málaga se ha realizado aplicando los test indicados a continuación.

El primer test aplicado, el Test de Rango fijo, consta de dos comprobaciones. Como resultado de la primera comprobación se obtiene que, efectivamente, no existe ningún dato de la serie de máximos anuales en ninguna de las setenta y dos estaciones meteorológicas que sea negativo, es decir, inferior a 0 mm, o superior a 508 mm de precipitación, por tanto, la primera comprobación da como resultado final cero outliers. La Tabla 3.5 muestra los resultados u outliers obtenidos de la segunda comprobación, es decir, el número de valores de precipitación máxima diaria de cada estación que superan a las efemérides tomadas como referencia de la AEMET (Tabla 3.4). La estación con mayor número de datos marcados por este test, es decir, registros que superaron los

umbrales máximos establecidos como efemérides, fue Benaoján CP, con 54 valores. Igualmente, la estación con menor número de datos marcados por este test fue Antequera el Águila, con tan sólo 2 valores que superan el umbral marcado por las efemérides.

Tabla 3.4. Efemérides para la estación de Málaga Aeropuerto (AEMET, 2012)

Mes	Precipitación (mm)	Día
Enero	121,5	18/01/1979
Febrero	151,0	22/02/1969
Marzo	102,1	27/03/2004
Abril	47,4	9/04/1994
Mayo	44,0	4/05/1946
Junio	72,5	14/06/1974
Julio	22,2	19/07/1942
Agosto	70,3	28/08/1987
Septiembre	313,0	27/09/1957
Octubre	109,4	31/10/2008
Noviembre	164,0	28/11/1942
Diciembre	132,7	18/12/2010

Tabla 3.5. Número de valores que superan las efemérides de la AEMET de cada estación.

Estación	Nº	Estación	Nº
Agujero	7	Conde de Guadalhorce	5
Alcaucín Cortijo	8	Contaderas Forestal	9
Alcaucín Forestal	42	Corchado Central	13
Alfarnate	19	Cuevas Becerro	17
Alhaurín el Grande	6	El Burgo	7
Aljaima	9	Fuente Piedra	3
Almargen	5	Fuente Piedra Herriza	4

Almogía los Llanes	8	Gobantes Vivero	9
Álora	7	Humilladero	3
Alozaina	13	Hundidero Pto. Sapo	14
Alpandeire	27	Istán	29
Antequera Águila	2	Jimena Libar CE	19
Archidona	5	Las Mellizas	9
Arriate	4	Málaga Oficina	4
Benahavis	14	Marbella Inst. Laboral	11
Benalmádena	5	Mijas Faro Cal	9
Benamargosa	10	Moclinejo	9
Benamocarra	12	Montejaque CE	7
Benaolán CP	54	Nerja	5
Bobadilla Estación	7	Ojén	19
Borregos	5	Parauta Taramal	13
Buitreras CE	21	Parchite	9
Buitreras Presa	21	Peña Enamorados	5
Campillos	8	Periana	11
Canillas Aceituno	15	Pizarra	12
Cartajima	37	Rincón de la Victoria	4
Cártama Estación	6	Riogordo	8
Casabermeja	6	Ronda CE	4
Casabermeja VP	8	SP Alcántara	14
Casapalma	7	Tolox Millanas	5
Casarabonela Forestal	27	Torrox	7
Casares	25	Vegueta Grama	12
Chorro Estación	12	Vélez Málaga	7

Coín	20	Villanueva de Tapias	5
Colmenar	12	Viñuela	10
Cómpeta	20	Yedra	8

De forma análoga se ha aplicado el Test de Rango dinámico, en el que se han usado distintos valores de f , haciéndolo variar desde 0,2 hasta 5,0 en cada una de las estaciones meteorológicas del área de estudio. Se ha establecido así una relación entre el número de errores potenciales y el factor f empleado.

Por último se ha aplicado el Test de Persistencia, relacionando igualmente los errores potenciales con los distintos valores de f aplicados, que han oscilado entre 0,2 y 5, igual que en el Test de Rango Dinámico explicado anteriormente. El resultado obtenido de la aplicación de cada uno de los tests es un conjunto de datos potencialmente erróneos cuyo valor va disminuyendo a medida que f va aumentando.

Como toma de decisiones dentro de la validación de la serie de precipitación inicial se ha seleccionado, en primer lugar, un valor de f que marque el 1,5% de los datos (Estévez *et al.*, 2010), tanto para el Test de Rango dinámico como para el Test de Persistencia, obteniendo un gráfico particular para cada una de las estaciones estudiadas. En las Figuras 3.6 y 3.7 se representa dos estaciones como ejemplo de cada uno de los tests comentados.

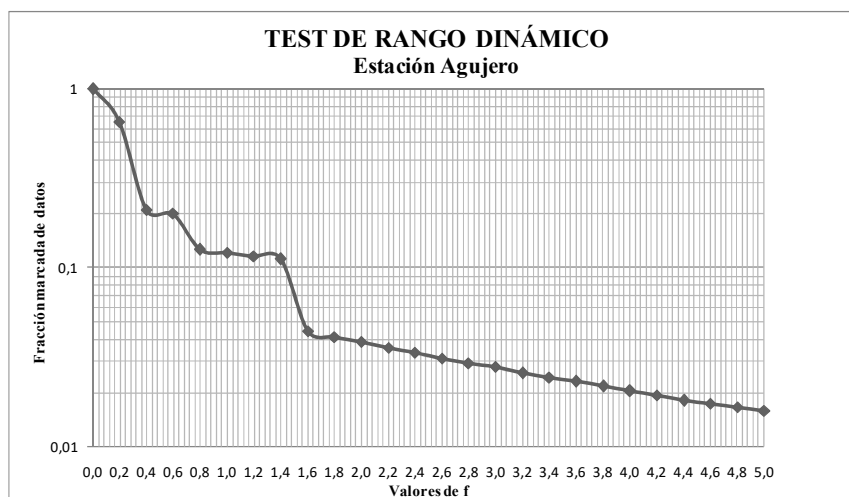


Figura 3.6. Fracción marcada de datos en la estación Agujero. Test de Rango Dinámico

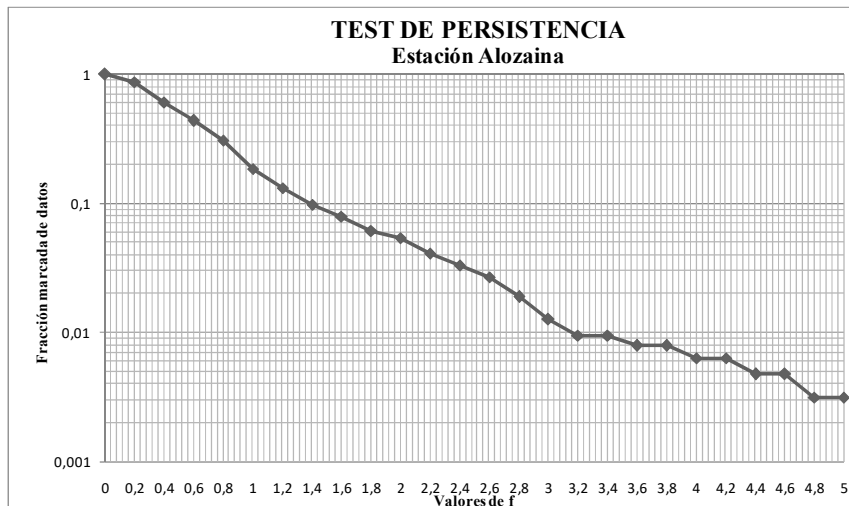


Figura 3.7. Fracción marcada de datos en la estación Alozaina. Test de Persistencia.

De esta primera etapa se han seleccionado los outliers (1.5% flagged) comunes en los datos marcados por ambos tests, es decir, por el Test de Rango Dinámico y por el Test de Persistencia.

Posteriormente se ha realizado la comprobación de que los datos marcados en ambos test, es decir, los que se encuentran dentro de ese 1,5% de datos eliminados, superan o no los valores de las efemérides registradas por la AEMET (Tabla 3.4), de manera que obtendremos un último conjunto de registros con aquellos datos marcados por los tres procedimientos descritos, es decir, el Test de Rango Dinámico, el Test de Persistencia y el Test de Rango Fijo donde se comparan los datos con las efemérides.

Continuando con el proceso de validación y en base a los resultados obtenidos en García-Marín *et al.*, (2011), se ha tenido en cuenta la diferencia entre aquellas estaciones que se encuentran en la misma región homogénea que Málaga Aeropuerto, tomada como estación de referencia por la presencia de datos en la AEMET, de aquellas que se encuentran fuera de ella, por tanto, situadas a más distancia estadística de Málaga, en las que las condiciones meteorológicas pueden oscilar notablemente. Según García-Marín *et al.*, (2011), las estaciones incluidas en la región homogénea en la que Málaga está incluida son Agujero, Aljaima, Almogia los Llanes, Alora, Benahavis,

Benalmádena, Buitreras CE, Cártama, Coín, Corchado Central, Málaga Oficina y Marbella Instituto Laboral.

Para la depuración final de la serie de precipitación inicial se ha tenido en cuenta, por tanto, el criterio de que en las estaciones cercanas a Málaga Aeropuerto y, por tanto, situadas en su misma región, se tendrán en cuenta los datos marcados como potencialmente erróneos por el Test de Rango Dinámico, el Test de Persistencia y el Test de Rango Fijo empleando las efemérides de la AEMET, obteniéndose un rango de datos marcados entre 1 y 7, en las estaciones de Málaga Oficina y Aljaima, respectivamente (Tabla 3.6). Para el caso de aquellas estaciones situadas fuera de la citada región los datos que se tendrán en cuenta serán los marcados por el Test de Rango Dinámico y el Test de Persistencia únicamente, en cuyo caso se han obtenido el mínimo en la estación de Borregos (8 datos marcados) y el máximo en Almargen y Cartajima (24 datos marcados), como se puede apreciar en la Tabla 3.6.

La citada Tabla muestra por lo tanto un resumen con el número de datos iniciales, proporcionados por la Confederación Hidrográfica de la Cuenca Sur, el número de datos iniciales de la serie depurada, resultado de aplicar el correspondiente procedimiento de validación, y el número de datos descartados de cada una de las setenta y dos estaciones meteorológicas.

Tabla 3.6. Número de datos iniciales, depurados y descartados.

ID	Estación	Iniciales	Depurados	Descartados
1	Agujero	17.532	17.530	2
2	Alcaucín Cortijo	19.358	19.342	16
3	Alcaucín Forestal	19.723	19.712	11
4	Alfarnate	21.549	21.537	12
5	Alhaurín el Grande	13.149	13.138	11
6	Aljaima	19.723	19.716	7
7	Almargen	20.088	20.064	24
8	Almogía los Llanes	18.627	18.623	4
9	Alora	19.723	19.719	4
10	Alozaina	20.454	20.433	21
11	Alpandeire	21.915	21.904	11
12	Antequera el Águila	20.819	20.802	17

13	Archidona	23.376	23.357	19
14	Arriate	19.358	19.348	10
15	Benahavis	17.532	17.529	3
16	Benalmádena	13.149	13.147	2
17	Benamargosa	21.184	21.163	21
18	Benamocarra	21.184	21.173	11
19	Benaoján CP	14.975	14.968	7
20	Bobadilla Estación	20.454	20.437	17
21	Borregos	18.627	18.619	8
22	Buitreras CE	21.549	21.545	4
23	Buitreras Presa	20.819	20.798	21
24	Campillos	20.088	20.075	13
25	Canillas Aceituno	21.184	21.176	8
26	Cartajima	20.819	20.795	24
27	Cartama Estación	18.993	18.990	3
28	Casabermeja	20.088	20.080	8
29	Casabermeja VP	19.358	19.345	13
30	Casapalma	14.975	14.970	5
31	Casarabonela Forestal	19.723	19.705	18
32	Casares	20.088	20.068	20
33	Chorro Estación	18.993	18.977	16
34	Coín	20.819	20.813	6
35	Colmenar	20.819	20.802	17
36	Cómpeta	21.184	21.173	11
37	Conde Guadalhorce	20.819	20.801	18
38	Contaderas Forestal	16.071	16.056	15
39	Corchado Central	21.184	21.180	4
40	Cuevas Becerro	20.088	20.075	13
41	El Burgo	20.819	20.797	22
42	Fuente Piedra	18.627	18.614	13
43	Fuente Piedra Herriza Herriza	13.149	13.137	12
44	Gobantes Vivero	20.088	20.073	15
45	Humilladero	19.723	19.708	15
46	Hundidero Pto. Sapo	16.801	16.785	16

47	Istán	21.184	21.165	19
48	Jimena Líbar CE	19.723	19.714	9
49	Las Mellizas	19.723	19.712	11
50	Málaga Oficina	14.244	14.243	1
51	Marbella Inst Laboral	19.723	19.719	4
52	Mijas Faro Cal	21.184	21.181	3
53	Moclinejo	21.915	21.904	11
54	Montejaque CE	21.184	21.165	19
55	Nerja	19.358	19.351	7
56	Ojén	18.993	18.988	5
57	Parauta Taramal	19.358	19.353	5
58	Parchite	19.723	19.710	13
59	Peña Enamorados	20.819	20.804	15
60	Periana	22.280	22.268	12
61	Pizarra	19.723	19.719	4
62	Rincón de la Victoria	13.149	13.145	4
63	Riogordo	20.088	20.074	14
64	Ronda CE	21.915	21.899	16
65	SP Alcántara	19.723	19.718	5
66	Tolox Millanas	13.149	13.146	3
67	Torrox	20.454	20.446	8
68	Vegueta Grama	20.819	20.806	13
69	Vélez Málaga	21.914	21.902	12
70	Villanueva Tapia	18.262	18.249	13
71	Viñuela	21.184	21.172	12
72	Yedra (La)	19.723	19.709	14

Para el caso concreto de la serie de datos horarios de precipitación para la estación “Málaga Aeropuerto” y después de aplicar el Test de Rango y el Test de Persistencia para la validación de los datos de lluvia, no se detectaron valores anómalos o sospechosos.

3.2.2.2. Resultados de validación en las estaciones de las regiones del Maule y Bio-Bio (Chile)

En el caso de las estaciones chilenas y siguiendo la metodología previamente descrita, el Test de Rango no detectó ningún error en los datos de precipitación diaria disponibles para la región del Maule. Sin embargo, el Test de Persistencia sí marcó algunos datos como sospechosos. Después de llevar a cabo una inspección manual de los mismos, solo se descartaron algunos datos de 13 de las estaciones disponibles (Tabla 3.7).

Tabla 3.7. Datos descartados en estaciones de la región del Maule (Chile)

ID	Nombre	Período	Test de Rango	Test de Persistencia
1	Ancoa Embalse	1957-2013	0	26
3	Armerillo	1948-2013	0	3
4	Bullileo Embalse	1930-2013	0	4
5	Colbún Maule Sur	1961-2013	0	3
13	El Manzano	1976-2013	0	3
23	Liguay	1975-2013	0	6
30	Parral	1964-2013	0	5
35	Río Ancoa Morro	1999-2013	0	3
36	Río Claro Rauquén	1999-2013	0	3
37	Río Loncomilla	2001-2013	0	3
39	Río Mataquito	2001-2013	0	4
46	San Manuel	1956-2013	0	4
50	Talca UC	1982-2013	0	3

En el caso de la serie de datos horarios de precipitación para la estación ‘Bullileo’, después de aplicar el Test de Rango y el Test de Persistencia para la validación de los datos de lluvia, no se detectaron valores anómalos o sospechosos.

Para los datos diarios disponibles en las estaciones de la región del Bio-Bio, ninguno de los tests aplicados marcó datos como sospechosos, por lo que las series originales serán la base del trabajo posterior.

4. APLICACIÓN Y RESULTADOS

4.1. Formación de Regiones Homogéneas en base a parámetros multifractales

El Análisis Regional de Frecuencias (ARF) de variables hidrometeorológicas es una alternativa al Análisis Local de Frecuencias (ALF) siempre que sea posible agrupar las series de datos disponibles en regiones homogéneas desde el punto de vista de la variable que se esté analizando. Como se verá en los apartados posteriores, la fase más compleja de este análisis regional es la formación de regiones homogéneas. Tradicionalmente se tiende a la utilización de las coordenadas geográficas, si bien no siempre resulta exitoso su uso pues la cercanía física no garantiza la similitud en la función de distribución de las series de datos analizadas.

Se propone en esta Tesis el uso de las características multifractales de las series de precipitación para formar regiones homogéneas. A lo largo de los apartados 4.1.2, 4.1.3 y 4.1.4 se describirá la formación de regiones en base a una serie de parámetros característicos obtenidos a partir de la aplicación de los algoritmos multifractales descritos en el apartado de Fundamentos Teóricos. Previamente se describen los aspectos fundamentales del ARF.

4.1.1. Análisis Regional de Frecuencias

El objetivo fundamental del ARF es relacionar la magnitud de extremos con su frecuencia de ocurrencia a través del uso de distribuciones de probabilidad (Ayuso, 2012), comunes a todas las estaciones consideradas en la región de estudio. Este análisis permite, por tanto, calcular datos para un determinado sitio de interés utilizando datos de otros lugares distintos a los del sitio en cuestión. Así, si se cuenta con N sitios, cada uno de ellos con n años de registros de eventos máximos, puede esperarse que $N \times n$ datos darán estimaciones más precisas de cuantiles tan extremos como Q_{Nn} (Hosking y Wallis, 1997).

En la práctica surgen algunos problemas ya que las distribuciones de frecuencias en lugares distintos no son idénticas y las series de registros pueden no ser estadísticamente independientes. Sin embargo un buen uso del ARF puede conducir a estimaciones

suficientemente precisas y de gran utilidad en diversos problemas reales (García-Marín, 2007).

Hosking y Wallis (1997), en su libro ‘Regional Frequency Analysis’ proponen el uso conjuntos del método del Índice de Avenida (Kite, 1977) y de los Momentos Lineales, Momentos-L, (del inglés *L-Moments*) para obtener el valor de los cuantiles de variables hidrológicas mediante Análisis Regional.

4.1.1.1 Método del Índice de Avenida

El término Índice de Avenida, aplicado por primera vez en el estudio de avenidas por Dalrymple en 1960, y de ahí su nombre, se ha utilizado desde entonces en el análisis de frecuencias de otras variables diferentes a la original, entre ellas las precipitaciones máximas (e.g. Sáenz de Ormijama *et al.*, 1991).

El método regional del Índice de Avenida se basa en el uso simultáneo de datos procedentes de un número de estaciones próximas que pueden considerarse homogéneas, de forma que los datos distribuidos en el espacio sustituyen en cierto modo a los datos distribuidos en el tiempo, que son bastante limitados (Sáenz de Ormijama *et al.*, 1991). Se obtiene así una relación entre los cuantiles sin dimensiones o índice de avenida (relaciones entre la estimación de la variable asociada a un cierto periodo de retorno y la media anual) y el periodo de retorno (T), que se conoce con el nombre de curva regional de frecuencia (García-Marín, 2007).

Según Hosking y Wallis (1997), si los datos de un conjunto de estaciones en una región poseen distribuciones idénticas, excepto por un factor de escala (el índice de avenida) que refleja las características de la variable en cada estación, conocidos los cuantiles a escala regional y los valores medios locales, se pueden extrapolar extremos locales con cierta fiabilidad.

Así, si se dispone de datos en N estaciones, cada estación i con un tamaño de muestra n_i , y Q_{ij} , $j=1 \dots n_i$ datos observados, $Q_i(F)$, $0 < F < 1$ será la función cuantil de la distribución de frecuencia de cada lugar. Puede escribirse entonces:

$$Q_i(F) = \mu_i q(F), i = 1 \dots N \quad (4.1)$$

donde μ_i es el índice de avenida, que adoptará el valor de la media de la distribución en la estación considerada y $q(F)$, que es la curva regional de frecuencias (función cuantil adimensional común a todas las estaciones de la región).

La base para estimar la curva regional de frecuencias es el cálculo de $q_{ij} = Q_{ij} / \mu_i$, $j = 1 \dots n_i$ e $i = 1 \dots N$, y será función de un conjunto de parámetros indeterminados $(g_1 \dots g_p)$ a los que se les pueden asignar momentos de grado superior de la función como el coeficiente de variación, el de sesgo o incluso determinados cocientes de L-momentos. Por tanto, la curva regional también se puede expresar como $q(F, g_1 \dots g_p)$.

Una vez calculados dichos parámetros para cada una de las estaciones se combinarán para dar una estimación regional, a modo de media ponderada, de manera que la estación i tendrá un peso proporcional a n_i , es decir:

$$g_K^R = \sum_{i=1}^N n_i g_K^{(i)} / \sum_{i=1}^N n_i \quad (4.2)$$

Sustituyendo los parámetros medios calculados en $q(F)$ se obtiene una estimación de la curva regional de frecuencia:

$$q(F) = q(F; g_1^R \dots g_K^R) \quad (4.3)$$

Una vez obtenidos tanto el estimador μ_i del índice de avenida como el estimador $q(F)$ de la curva regional de frecuencia, la estimación de los cuantiles de la estación i se reduce a la aplicación de la ecuación de partida (Ecuación 4.1).

Conviene señalar que el método del índice de avenida parte de asumir una serie de supuestos:

- Las observaciones de una estación están idénticamente distribuidas.
- Las observaciones en una estación son independientes.
- Observaciones en diferentes estaciones son independientes.
- Las distribuciones de frecuencia en distintos sitios de una misma región son idénticas excepto por un factor de escala.
- La forma de la curva regional de frecuencias se especifica correctamente desde el punto de vista matemático.

Tal como explican Hosking y Wallis (1997), las dos primeras hipótesis se cumplen para la mayor parte de los casos ya que tanto los valores máximos como los acumulados anuales están, por lo general, libres de variaciones temporales. Además, es evidente que todos los métodos de análisis de frecuencias tienen que asumir que los eventos observados en el pasado constituyen la primera pista sobre lo que se puede esperar en el futuro. Sin embargo, la tercera hipótesis rara vez se cumple estrictamente, pues suele existir cierta correlación entre las observaciones de estaciones cercanas entre sí. Las dos últimas podrán ser alcanzadas de forma aproximada mediante una selección cuidadosa de las estaciones de cada región y la elección de distribuciones de frecuencia constantes con las observaciones (Erro Eslava, 2011).

4.1.1.2. Momentos Lineales

Los momentos lineales constituyen un sistema alternativo al método tradicional de los momentos convencionales para describir la forma de las funciones de distribución y proporcionar así medidas de localización, dispersión, sesgo, curtosis y otros aspectos de las distribuciones de probabilidad o de las muestras de datos (Ayuso, 2012). Son combinaciones lineales de los momentos ponderados probabilísticamente, introducidos por Greenwood *et al.*, (1979), cuya variable aleatoria X y función de distribución $F(.)$ quedan definidas por la expresión (4.4), Hosking y Wallis (1997),

$$M_{p,r,s} = E[X^p \{F(X)\}^r \{1-F(X)\}^s] \quad (4.4)$$

donde E indica el valor de la Esperanza y p , r y s son números reales. Son particularmente útiles los casos en que $p=1$, $r=0$ y $s=0$, denotados como $\alpha_r = M_{1,0,r}$ y $\beta_r = M_{1,r,0}$ respectivamente, siendo $r=0, 1, 2, \dots$

En términos de momentos de probabilidad ponderada, los cuatro primeros momentos lineales λ_i quedan definidos por las siguientes expresiones:

$$\lambda_1 = \alpha_0 = \beta_0 \quad (4.5)$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0 \quad (4.6)$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0 \quad (4.7)$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (4.8)$$

El momento lineal de primer orden (λ_1) es el parámetro de localización o media de la distribución; el de segundo orden (λ_2) mide la escala de distribución, indicando el grado de dispersión de los datos; el de tercer orden (λ_3) hace referencia a la asimetría y el de cuarto orden (λ_4) indica la curtosis o apuntamiento de la distribución.

No obstante, para poder comparar distintas estaciones es imprescindible definir los momentos lineales de forma adimensional, para lo que se calculan los ratios o cocientes de momentos lineales, dividiendo aquellos de mayor orden por la medida de escala λ_2 (Hosking y Wallis, 1997):

$$\tau_r = \lambda_r / \lambda_2, r = 3, 4, \dots \quad (4.9)$$

lo que dará una idea de la forma de la distribución con independencia de su escala de medida. Adquieren particular importancia los coeficientes lineales de variación (LC_v), de asimetría o sesgo (LC_s) y de curtosis (LC_k):

$$LC_v = \lambda_2 / \lambda_1 \quad (4.10)$$

$$LC_s = \lambda_3 / \lambda_2 \quad (4.11)$$

$$LC_k = \lambda_4 / \lambda_2 \quad (4.12)$$

Se puede afirmar que, conceptualmente, los L-momentos son semejantes a los momentos convencionales. Sin embargo Hosking y Wallis (1997) constatan que son varias las ventajas de los primeros desde el punto de vista teórico: capacidad de caracterizar una amplia gama de funciones de distribución, mayor robustez en presencia de valores anómalos (outliers) y menor susceptibilidad a la asimetría o al sesgo que los momentos convencionales.

A nivel práctico, los L-momentos con los que se suele trabajar son la media λ_1 , la escala λ_2 , el coeficiente de variación (LC_v) τ , el coeficiente de asimetría o sesgo (LC_s) τ_3 y el coeficiente de curtosis (LC_k) τ_4 , expresados como l_1 , l_2 , t , t_3 y t_4 cuando hacen referencia a valores muestrales en las observaciones.

4.1.1.3. Pasos a seguir en el Análisis Regional de Frecuencias

El Análisis Regional de frecuencias consta básicamente de cinco etapas, y en cada una de ellas se incluye el cálculo de una serie de estadísticos cuyos valores marcarán las decisiones a tomar en cada uno de los pasos:

— *Análisis y filtrado primario de las observaciones.*

Este es el primer paso en cualquier tipo de análisis estadístico, encaminado a detectar posibles errores de los datos de partida que hayan podido afectar a la distribución de frecuencias en alguna de las estaciones de estudio.

— *Identificación de regiones homogéneas.*

Consiste en agrupar los distintos lugares en regiones. Las estaciones no podrán incluirse en regiones en las que el criterio de homogeneidad no se cumpla estrictamente, aunque una homogeneidad aproximada sería suficiente para garantizar una precisión superior del análisis regional sobre el análisis local. Calculando diversos estadísticos del lugar y comparando la variabilidad entre las estaciones, así como los valores esperados para la región, puede estudiarse el grado de homogeneidad de la misma.

— *Elección de la distribución de frecuencia para la región.*

Consiste en seleccionar la distribución regional de frecuencia $q(F)$ más apropiada para cada región mediante pruebas de bondad de ajuste a las diversas funciones posibles. Es necesario determinar el grado de similitud entre los estadísticos referentes a los datos observados y los que se esperarían si éstos procedieran de una muestra aleatoria extraída de una población que obedeciese a una función de distribución de frecuencias conocida.

— *Estimación de los cuantiles regionales de frecuencia.*

Una vez elegida la función de distribución adecuada para la región, se estiman los cuantiles asociados a un cierto periodo de retorno. Para ello la distribución se ajustará separadamente en cada una de las estaciones, combinándose a posteriori con el fin de obtener una media regional. Se seguirá el método del índice de avenida.

— *Análisis de la robustez de los modelos estadísticos seleccionados.*

Para que los resultados alcancen el mayor grado de utilidad práctica posible y así poder calibrar su fiabilidad se evalúa la magnitud de la incertidumbre mediante la simulación de Monte Carlo, lo que permitirá comprobar si los cuantiles estimados mediante ARF se encuentran dentro de un determinado intervalo de confianza previamente definido.

4.1.1.4. Discordancia

La primera prueba estadística a realizar en un análisis regional de frecuencias es la medida de la discordancia D_i , estadístico que compara los L-momentos de cada estación en una región con los del resto de estaciones y los valores medios de cada región (Hosking y Wallis, 1997).

El objetivo de la discordancia es, por tanto, determinar aquellas estaciones que son discordantes con el resto del grupo. Para ello se define el vector de L-momentos lineales (LC_v , LC_s , LC_k) de cada estación y, considerándolo como un punto en un espacio tridimensional, se considerará discordante cualquiera de ellos que quede lejos del centro de la nube de puntos formada por el resto de estaciones.

Numéricamente la medida de la discordancia D_i del sitio i se define a partir de un estadístico:

$$D_i = \frac{1}{3} N(u_i - \bar{u})^T A^{-1} (u_i - \bar{u}) \quad (4.13)$$

siendo $u_i = [t^{(i)} t_3^{(i)} t_4^{(i)}]^T$ el vector que contiene los valores de los ratios de L-momentos t , t_3 y t_4 de la estación i , con el superíndice T indicando la transposición del mismo; $\bar{u} = N^{-1} \sum_{i=1}^N u_i$ es la media no ponderada del grupo; $A = \sum_{i=1}^N (u_i - \bar{u})(u_i - \bar{u})^T$ es la matriz de la suma de cuadrados y productos cruzados y N el número de estaciones.

La estación i se considera discordante si el valor de D_i es elevado y ello está condicionado por el número de estaciones del grupo. Hosking y Wallis (1997)

establecen que una estación ha de ser catalogada como discordante si su D_i supera un valor crítico establecido en función del número de sitios en la región (Tabla 4.1).

Tabla 4.1. Valores críticos de la discordancia (Hosking y Wallis, 1997)

Nº Estaciones	Valor crítico D_i
5	1,333
6	1,648
7	1,917
8	2,140
9	2,329
10	2,491
11	2,632
12	2,757
13	2,869
14	2,971
≥ 15	3,000

A nivel práctico Hosking y Wallis (1997) recomiendan el empleo de la medida de la discordancia en dos ocasiones, en primer lugar al inicio del análisis, procurando abarcar la totalidad de estaciones o regiones con el fin de que el número de ellas sea elevado. Con esta primera prueba se trata de detectar aquellos sitios que se diferencian significativamente del resto. Posteriormente, una vez identificadas las regiones homogéneas, es aconsejable realizar una segunda prueba o medida de la discordancia en cada región y aquellos sitios que resulten discordantes se tendrán en cuenta a la hora de valorar la posibilidad de trasladarlos a otra región, de hecho son habituales los eventos que afectan únicamente a algunas estaciones dentro de una región homogénea.

4.1.1.5. Homogeneidad

De todas las etapas del ARF, la identificación de regiones homogéneas es normalmente, la más complicada ya que requiere de la toma de decisiones y aporta un gran valor subjetivo. El objetivo que se plantea es, por tanto, la formación de grupos de estaciones que satisfagan aproximadamente la condición de homogeneidad. Para ello suele realizarse una prueba de heterogeneidad H , cuyo estadístico evalúa cuándo un grupo de estaciones pueden ser tratadas en conjunto como una región homogénea. La prueba utilizada compara la variabilidad de los L- momentos muestrales de las estaciones con la esperada tras emplear técnicas de simulación, utilizando para ello la varianza ponderada del coeficiente de variación lineal de cada estación según el estadístico V ,

$$V = \left\{ \sum_{i=1}^N n_i (t^{(i)} - t^R)^2 / \sum_{i=1}^N n_i \right\}^{1/2} \quad (4.14)$$

siendo $t_R = \sum_{i=1}^N n_i t^{(i)} / \sum_{i=1}^N n_i$, N el número de estaciones, n_i el número de registros y $t^{(i)}$ el valor de los momentos lineales de la muestra.

Se trata de calcular la media y la desviación estándar de V , es decir, μ_v y σ_v , para una serie de simulaciones N_{sim} , cuyo número se seleccionará de manera que las estimaciones μ_v y σ_v sean fiables, teniendo en cuenta que Hosking y Wallis (1997) consideran que N_{sim} igual a 50 es un valor adecuado.

Una vez calculado V , la medida de la heterogeneidad vendrá dada por:

$$H = \frac{(V - \mu_v)}{\sigma_v} \quad (4.15)$$

Una región se considerará aceptablemente homogénea si $H < 1$, posiblemente heterogénea si $1 \leq H < 2$ y definitivamente heterogénea para valores de $H > 2$.

4.1.1.6. Ajuste a una función de distribución de frecuencias

Una vez alcanzada la homogeneidad de una región ha de encontrarse la función de distribución más apropiada para ella. Generalmente, se puede decir que no existe una homogeneidad total, de ahí que no exista una única distribución posible. Por tanto, el

objetivo se centra en encontrar la función de distribución que proporcione las estimaciones más precisas para cada estación dentro de la región de estudio.

Es importante destacar que la distribución elegida no tiene por qué ser la que más se aproxime a las observaciones. El hecho de que exista una función a la que los datos se ajusten bien no garantiza que los valores futuros vayan a ser coherentes con los del pasado. En este sentido es preferible anteponer la elección de un modelo robusto en estimaciones a la selección de la distribución que proporcione un mejor ajuste (Erro Eslava, 2011). Por lo general, en el estudio de eventos extremos de precipitaciones, avenidas o sequías, son de particular relevancia las estimaciones que recaen en la cola de la distribución de frecuencia (Erro Eslava, 2011).

Para el análisis regional de frecuencias, Hosking y Wallis (1997) proponen la elección de una función de distribución, entre las siguientes: Logística Generalizada (*GLO*), General de Valores Extremos (*GEV*), General de Pareto (*GPA*), Log Normal (*LN3*) y Pearson Tipo III (*PE3*) (no obstante, el abanico de funciones de distribución puede ampliarse sin problema). Hosking y Wallis (1997) justifican la elección entre estas distribuciones por el hecho de poseer tres parámetros. Si son varias las funciones que se ajustan correctamente se opta por la más robusta, es decir, la que esté más capacitada para proporcionar buenas estimaciones aun cuando los valores futuros procedan de una distribución diferente a la ajustada.

Son muchas las alternativas disponibles para evaluar la bondad del ajuste de una muestra de observaciones a una determinada función de distribución. Sin embargo, en el contexto del ARF, Hosking y Wallis (1997) proponen trabajar directamente sobre los L-momentos regionales. Definen, además, una medida de la bondad de ajuste para evaluar hasta qué punto los datos se aproximan a una distribución concreta y seleccionar así el mejor ajuste.

La bondad del ajuste juzgará en qué medida los coeficientes lineales de sesgo y curtosis de la distribución seleccionada se ajustan bien al promedio regional de tales coeficientes con los datos observados. Para medir la bondad del ajuste Hosking y Wallis (1997) establecen el estadístico,

$$Z^{DIST} = (\tau_4^{DIST} - t_4^R + B_4) / \sigma_4 \quad (4.16)$$

siendo τ_4^{DIST} el coeficiente lineal de sesgo, t_4^R el coeficiente lineal de curtosis, σ_4 la desviación estándar de t_4^R y B_4 un coeficiente, función del número de simulaciones, para mejorar el ajuste de t_4^R .

Se considera que el ajuste a la distribución es adecuado si el estadístico Z^{DIST} es suficientemente cercano a cero, siendo un valor razonable para este criterio un grado de significación del 90%, que corresponde a un valor absoluto de Z^{DIST} menor o igual a 1,64 (Hosking y Wallis, 1997).

$$|Z^{DIST}| \leq 1,64 \quad (4.17)$$

Para todas aquellas regiones que cumplan la condición $|Z^{DIST}| \leq 1,64$ se calcula su curva regional de frecuencias. Si éstas son aproximadamente iguales cualquiera de las distribuciones es adecuada, si bien interesa optar por el método más robusto frente a posibles errores cometidos en la definición de la región. En caso de que las curvas difieran, puede existir un problema de escasez de datos que conduce a diferencias operacionales importantes. Para estas situaciones, la robustez adquiere especial trascendencia y el empleo de funciones tipo Kappa (de cuatro parámetros) o Wakeby (de cinco parámetros) podría ser aconsejable (Erro Eslava, 2011).

Puede ocurrir que ninguna de las distribuciones utilizadas cumpla el criterio $|Z^{DIST}| \leq 1,64$, siendo la presencia de un número elevado de estaciones por región la causa más probable. Si esto ocurre, una solución podría ser aumentar el número de funciones de distribución. La formación de subregiones homogéneas puede ser otra opción.

En regiones que no son aceptablemente homogéneas, no existe razón para dar por supuesto que una función de distribución pueda caracterizarlas adecuadamente, pero tampoco para descartar su existencia.

Para regiones claramente heterogéneas Hosking y Wallis (1997) recomiendan anteponer los L-momentos locales a los regionales de cara a discriminar entre las distribuciones candidatas. Es evidente que, ante la heterogeneidad, basar el ajuste en los valores

medios regionales carece de sentido, pues no son representativos de cada estación de la región.

4.1.1.7. Métodos para formar regiones

Análisis de Componentes Principales (ACP)

A la hora de buscar información sobre una muestra de datos, lo más frecuente es tomar el mayor número posible de variables. Sin embargo, esto hace difícil visualizar relaciones entre las mismas. Otro problema que se plantea es la fuerte correlación que muchas veces se presenta entre las variables, es decir, si tomamos demasiadas variables lo normal es que estén relacionadas o que midan lo mismo bajo distintos puntos de vista.

Se hace necesario, pues, reducir el número de variables. Es importante resaltar el hecho de que el concepto de mayor información se relaciona con el de mayor variabilidad, es decir, cuanto mayor sea la variabilidad de los datos (varianza) se considera que existe mayor información, lo que está relacionado con el concepto de entropía.

La técnica del ACP fue desarrollada por Pearson a finales del siglo XIX y posteriormente fue estudiada por Hotelling en los años 30 del pasado siglo.

Para estudiar las relaciones que se presentan entre p variables correlacionadas se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí, es decir, sin repetición o redundancia de información, llamado conjunto de componentes principales.

Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra.

De modo ideal, se buscan $m < p$ variables que son combinaciones lineales de las p originales y que estén incorreladas, recogiendo la mayor parte de la información o variabilidad de los datos. Si las variables originales están incorreladas de partida, entonces no tiene sentido realizar este análisis.

El objetivo de este análisis es, pues, la identificación de los componentes principales, es decir, averiguar qué información de la muestra aportan. Sin embargo este es un

problema difícil que a menudo resulta bastante subjetivo. Habitualmente, se conservan sólo aquellos componentes que recogen la mayor parte de la variabilidad, hecho que permite representar los datos según dos o tres dimensiones.

En el caso de la lluvia, Hosking y Wallis (1997) aconsejan la formación de regiones homogéneas teniendo en cuenta las características del sitio, como la localización, la altura o el valor de la precipitación. Con características de este tipo, podría realizarse un ACP que permita la formación de regiones. Cada una de las estaciones analizadas podría caracterizarse por parámetros representativos, como por ejemplo, longitud (coordenada x , asociada con la procedencia del frente de lluvia), latitud (coordenada y , que representa la proximidad al mar y el efecto regulador de la lluvia), altitud (coordenada z , que proporciona la altura sobre el nivel del mar y da una idea de la orografía de la zona) y precipitación máxima para una duración y un periodo de retorno determinado.

El uso del ACP posibilita la formación de regiones homogéneas con aquellas estaciones meteorológicas que presenten características comunes con los parámetros seleccionados anteriormente (García-Marín *et al.*, 2011).

Análisis Clúster

Diversos autores han establecido criterios alternativos de agrupación de estaciones, entre ellos la delimitación según áreas administrativas (Beable y Mckerchan, 1982), demasiado arbitraria y carente de justificación física o, a pequeña escala, la inspección subjetiva de las características físicas de las estaciones (Shafer, 1990). No obstante Hosking y Wallis (1997) consideran el análisis clúster como una de las alternativas más apropiadas.

El análisis clúster es un método estándar de análisis multivariante por el que se asocia un vector a cada una de las estaciones, que son divididas y agrupadas atendiendo precisamente a la similitud entre vectores.

En el contexto del ARF hay que tener en cuenta algunas consideraciones. Los clústers se forman por agrupación de lugares cuyas características físicas son parecidas y la mayoría de los algoritmos miden la similitud de dichas características atendiendo a las distancias euclídeas en el espacio.

Es necesario destacar, igualmente, que no existe un número de clústers correcto sino que es la propia área la que marca la necesidad de agrupar en más o menos regiones. En este sentido, es evidente que los clústers con pocas estaciones corren el riesgo de mejorar muy poco la precisión de las estimaciones con respecto a los métodos del análisis local mientras que las regiones de gran tamaño pueden vulnerar el criterio de homogeneidad.

Se destaca además otro aspecto importante y es que el resultado que arroja un análisis clúster no tiene por qué ser el definitivo, ya que son muchos los ajustes de tipo subjetivo que se pueden emplear para mejorar la coherencia del reparto e incluso reducir la heterogeneidad de ciertas regiones, entre ellos, el movimiento de estaciones de una región a otra, la eliminación de ciertas estaciones, la subdivisión de regiones resultantes, la rotura de regiones mediante el traslado de estaciones de un clúster a otro y la combinación de regiones entre sí.

Existen diversos procedimientos para realizar un análisis clúster, entre los que destaca el método ‘K-means clustering’. Este método es un algoritmo que tiene como objetivo la partición de n observaciones en k grupos ($k < n$), en el que cada observación pertenece al grupo más cercano con la media. Los puntos se ajustan iterativamente, normalmente se aplican 10 iteraciones en el proceso, de manera que cada uno de los N puntos se asigna a cada una de las agrupaciones k y cada uno de los k grupos es la media de los puntos asignados (Bishop, 1995).

4.1.2. Formación de regiones homogéneas de precipitación máxima diaria anual en base a los valores de los parámetros multifractales q_{crit} (momento crítico) y γ_{max}

Para el desarrollo de esta primera aplicación de la agrupación en base a parámetros multifractales, se dispone de las series de precipitaciones máximas diarias de setenta y dos estaciones meteorológicas situadas en la provincia de Málaga, suministradas por la Confederación Hidrográfica de la Cuenca Sur (Véase apartado 3.1.1).

A partir de estos datos de partida y tras la correspondiente etapa de validación de datos (véase apartado 3.2) en la que se han aplicado el Test de Rango fijo, con sus dos comprobaciones pertinentes, el Test de Rango dinámico y el Test de Persistencia, se obtienen las series de precipitaciones máximas diarias anuales.

Con estas series, se inicia el Análisis Regional de Frecuencias (ARF). En primer lugar, se calculan los cinco primeros momentos lineales así como los coeficientes lineales de variación, sesgo y curtosis para todas las series de precipitación máxima existentes (Tabla 4.2).

Tabla 4.2. Valores de los L-momentos de las estaciones de la provincia de Málaga.

ID	Estación	$L_1 (\mu)$	$L_2 (\sigma)$	$L_2/L_1 (C_v)$	$L_3 (C_s)$	$L_4 (C_k)$	$L_5 (\lambda_5)$
1	Agujero	62,7854	13,6250	0,2170	0,2343	0,1806	0,0759
2	Alcaucín Cortijo	59,3051	8,7587	0,1477	0,1360	0,0912	0,0526
3	Alcaucín Forestal	93,6674	18,4514	0,1970	0,0028	0,1403	0,0926
4	Alfarnate	77,4333	11,7520	0,1518	-0,0308	0,0231	-0,0255
5	Alhaurín Grande	81,6000	14,5310	0,1781	0,2340	0,1230	-0,0667
6	Aljaima	55,9463	11,2361	0,2008	0,1135	0,0994	-0,0157
7	Almargen	43,4273	7,5430	0,1737	0,1643	0,1710	0,0079
8	Almogía los Llan	68,4265	17,0539	0,2492	0,2071	0,1023	0,0599
9	Alora	58,5389	12,1340	0,2073	0,0794	0,0087	-0,0299
10	Alozaina	71,1231	19,9213	0,2801	0,3198	0,1661	0,0816
11	Alpandeire	83,8544	12,8432	0,1532	-0,0104	0,1663	0,0873
12	Antequera Águil	44,5456	6,9781	0,1567	0,1138	0,0692	0,0136
13	Archidona	42,1828	7,2148	0,1710	0,1643	0,0877	-0,0533
14	Arriate	51,2038	10,1968	0,1991	0,3065	0,1800	0,0499
15	Benahavis	75,1694	16,4836	0,2193	0,2015	0,1562	0,0495
16	Benalmádena	72,5885	13,8069	0,1902	0,2036	0,1544	0,0538
17	Benamargosa	62,8000	11,0840	0,1765	0,1766	0,1990	0,0564
18	Benamocarra	61,4241	11,9513	0,1946	0,1704	0,1195	0,0516
19	Benaolán CP	115,3927	24,2311	0,2100	0,1951	0,1734	0,0920
20	Bobadilla Estac	43,0927	9,2073	0,2137	0,2129	0,1275	0,0406
21	Borregos	41,0471	7,9085	0,1927	0,1414	0,0770	-0,0108

ID	Estación	$L_1 (\mu)$	$L_2 (\sigma)$	$L_2/L_1 (C_v)$	$L_3 (C_s)$	$L_4 (C_k)$	$L_5 (\lambda_5)$
22	Buitreras CE	82,6879	16,3067	0,1972	0,1284	0,1254	0,0255
23	Buitreras Presa	87,8786	15,3038	0,1741	0,1106	0,1269	0,0185
24	Campillos	45,3811	7,7692	0,1712	0,2551	0,2320	0,1521
25	Canillas Aceituno	60,3923	9,4139	0,1559	0,1461	0,0677	0,0347
26	Cartajima	84,4000	16,1768	0,1917	0,1026	0,0557	0,0115
27	Cartama Estación	66,3932	14,1189	0,2127	0,1246	0,1175	0,0141
28	Casabermeja	56,0721	11,2842	0,2012	0,1835	0,1001	0,0153
29	Casabermeja VP	58,9942	12,3159	0,2088	0,0942	0,0241	0,0086
30	Casapalma	60,6528	12,0136	0,1981	0,0714	0,0555	-0,0276
31	Casarabonela	85,9114	18,8895	0,2199	0,0530	0,0940	0,0018
32	Casares	82,2455	16,6616	0,2026	0,1282	0,0965	0,0344
33	Chorro Estación	54,1373	11,7717	0,2174	0,2401	0,1670	0,0517
34	Coin	80,5140	18,8627	0,2343	0,1997	0,0751	-0,0132
35	Colmenar	65,0170	10,6912	0,1644	0,0614	0,0182	-0,0411
36	Cómpeta	70,8929	10,9977	0,1551	0,0805	0,1136	0,0808
37	Conde Guadalhor	45,6842	9,0576	0,1983	0,1923	0,1131	0,0479
38	Contaderas Fores	65,2977	12,7723	0,1956	0,2047	0,1471	0,0591
39	Corchado Central	74,6966	14,8888	0,1993	0,0670	0,1356	0,0079
40	Cuevas Becerro	63,4981	12,4648	0,1963	0,2274	0,1982	0,0573
41	El Burgo	57,0963	12,3224	0,2158	0,2903	0,1633	0,0630
42	Fuente Piedra	42,7745	7,9020	0,1847	0,1686	0,1760	0,0724
43	Fuente P.Herriza	40,4528	6,6353	0,1640	0,2476	0,2639	0,1670
44	Gobantes Vivero	49,9923	10,9312	0,2187	0,2456	0,1085	0,0916
45	Humilladero	38,7094	6,2537	0,1616	0,1699	0,2070	0,0970
46	Hundidero	72,8386	15,6634	0,2150	0,1269	0,1521	0,0342
47	Istán	82,4379	15,8799	0,1926	0,2253	0,1286	0,0403
48	Jimena Líbar CE	80,2962	13,4214	0,1671	0,1648	0,1871	0,1010

ID	Estación	$L_1 (\mu)$	$L_2 (\sigma)$	$L_2/L_1 (C_v)$	$L_3 (C_s)$	$L_4 (C_k)$	$L_5 (\lambda_5)$
49	Las Mellizas	58,0420	12,5613	0,2164	0,1417	0,1201	0,994
50	Málaga Oficina	63,7564	13,6116	0,2135	0,2581	0,2271	0,0935
51	Marbella Inst.	71,0660	18,5774	0,2614	0,2168	0,1094	0,0380
52	Mijas Faro Cal	54,0085	11,7503	0,2176	0,2641	0,2423	0,1425
53	Moclinejo	55,2281	10,2454	0,1855	0,0484	0,1074	-0,0061
54	Montejaque	57,3105	11,1823	0,1951	0,1702	0,1131	0,0564
55	Nerja	48,2596	9,1372	0,1893	0,1295	0,1130	0,0050
56	Ojén	91,1739	20,3198	0,2229	0,1135	0,0853	-0,0054
57	Parauta Taramal	81,9615	13,7981	0,1683	0,1092	0,0250	0,0338
58	Parchite	51,1538	10,4337	0,2040	0,3012	0,1664	0,0509
59	Peña Enamorado	43,8737	8,9032	0,2029	0,1880	0,1572	0,0310
60	Periana	58,3400	9,1002	0,1560	0,1412	0,1291	-0,0140
61	Pizarra	63,1462	15,0854	0,2389	0,1658	0,1294	0,0278
62	Rincón Victoria	51,5697	12,0049	0,2328	0,3228	0,3796	0,2082
63	Riogordo	57,9679	11,8773	0,2049	0,0468	0,1259	0,0532
64	Ronda CE	48,6167	8,0916	0,1664	0,0598	0,0700	0,0212
65	SP Alcántara	73,7040	14,9043	0,2022	0,2343	0,1237	0,0245
66	Tolox Millanas	75,7971	16,1634	0,2132	0,1659	0,1682	0,0186
67	Torrox	59,0455	10,9301	0,1851	0,0492	0,0637	0,0253
68	Vegueta Grama	67,9170	13,4393	0,1979	0,1336	0,0773	0,0420
69	Vélez Málaga	54,8630	9,0594	0,1651	0,2046	0,2851	0,1142
70	Villanueva Tapia	50,3816	7,4390	0,1477	0,1711	0,1969	0,0846
71	Viñuela	59,8482	10,6651	0,1782	0,2465	0,0975	-0,0014
72	Yedra	56,2722	9,1778	0,1631	-0,0683	0,0900	0,0527

Una vez obtenidos los L-Momentos y siguiendo con la metodología del ARF, se obtiene el valor del estadístico D_i (ecuación 4.13) para cada una de las setenta y dos estaciones de la primera región potencial, a la que llamaremos Málaga. Para regiones de más de 15 estaciones, el valor crítico de la discordancia es de 3,0 (Tabla 4.1). Considerando este límite, cuatro estaciones de las setenta y dos iniciales resultan discordantes y han de eliminarse del primer análisis: Alozaina ($D_i = 3,55$), Alpandeire ($D_i = 3,02$), Rincón ($D_i = 5,66$) y Yedra ($D_i = 3,06$). Con una nueva región de sesenta y ocho estaciones, se analiza nuevamente la discordancia, y aparecen dos nuevas estaciones discordantes, Alcaución Forestal y Vélez, con valores de D_i de 3,32 y 3,25, respectivamente. La nueva región cuenta ahora con sesenta y seis estaciones y ninguna de ellas como discordante. Se procede pues a la obtención del estadístico H (ecuación 4.15), para el que se obtiene un valor de 3,75, lo que pone de manifiesto la heterogeneidad de la región Málaga.

Se hace necesario pues la subdivisión de la región en subregiones a las que posteriormente se les aplicará el ARF para comprobar su homogeneidad. La división en subregiones se realizará en base a la caracterización multifractal de las series de datos y en concreto, a partir del valor del momento crítico q_{crit} obtenido mediante el método del escalado de momentos (véase apartado 2.2.1).

Para realizar la caracterización multifractal de los datos de lluvia en cada estación se procede al cálculo y representación del escalado de momentos. Esta representación muestra los q -ésimos momentos de intensidad de lluvia para escalas temporales que oscilan, en este caso, entre 1 día hasta aproximadamente 15 años, frente a la razón de escala (λ) para cada una de las estaciones (ecuación 2.5).

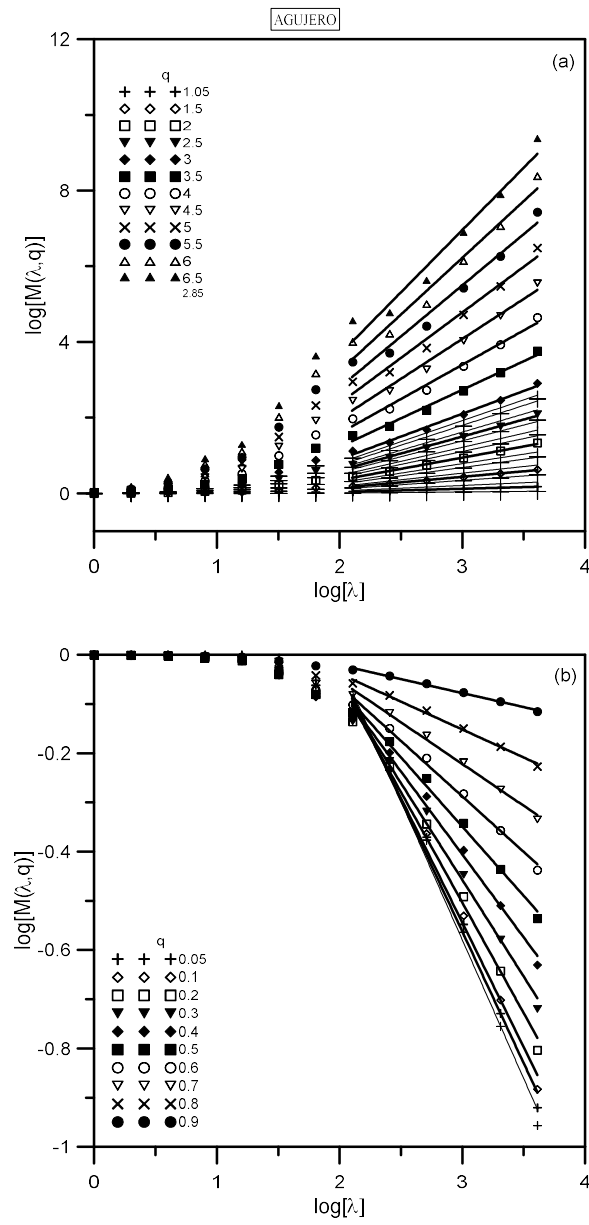


Figura 4.1. Representación log-log de los momentos q -ésimos. Estación Agujero.

A modo de ejemplo, se representa el escalado de momentos para la estación Agujero (Figura 4.1). Las líneas rectas de cada una de las partes (a) y (b) de la Figura 4.1 muestran que el comportamiento de escala para los momentos se produce desde 1 hasta 32 días. Para el resto de estaciones, los límites para el comportamiento de escala encontrado se muestran en la Tabla 4.3.

Tabla 4.3. Comportamiento de escala para los datos de las distintas estaciones

ID	Estación	Desde (días)	Hasta (días)
1	Agujero	1	32
2	Alcaucín Cortijo	1	32
3	Alcaucín Forestal	1	64
4	Alfarnate	1	32
5	Alhaurín El Grande	1	32
6	Aljaima	1	32
7	Almargen	1	32
8	Almogía los Llanes	1	16
9	Alora	1	16
10	Alozaina	1	32
11	Alpandeire	1	32
12	Antequera el Águila	1	16
13	Archidona	1	32
14	Arriate	1	32
15	Benahavis	1	32
16	Benalmádena	1	16
17	Benamargosa	1	16
18	Benamocarra	1	32
19	Benaolán CP	1	32
20	Bobadilla Estación	1	32

ID	Estación	Desde (días)	Hasta (días)
21	Borregos	1	32
22	Buitreras CE	1	16
23	Buitreras Presa	1	32
24	Campillos	1	32
25	Canillas Aceituno	1	32
26	Cartajima	1	64
27	Cartama Estación	1	32
28	Casabermeja	1	16
29	Casabermeja V.Pineda	1	16
30	Casapalma	1	16
31	Casarabonela Forestal	1	32
32	Casares	1	16
33	Chorro Estación	1	16
34	Coin	1	32
35	Colmenar	1	32
36	Cómpeta	1	32
37	Conde Guadalhor	1	32
38	Contaderas Forestal	1	16
39	Corchado Central	1	8
40	Cuevas Becerro	1	16
41	El Burgo	1	32
42	Fuente Piedra	1	16
43	Fuente Piedra Herriza	1	16
44	Gobantes Vivero	1	16
45	Humilladero	1	16
46	Hundidero Pto. Sapo	1	64
47	Istán	1	32

ID	Estación	Desde (días)	Hasta (días)
48	Jimena Líbar CE	1	16
49	Las Mellizas	1	16
50	Málaga Oficina	1	32
51	Marbella Inst. Laboral	1	16
52	Mijas Faro Cal	1	16
53	Moclinejo	1	32
54	Montejaque CE	1	16
55	Nerja	1	16
56	Ojén	1	32
57	Parauta Taramal	1	16
58	Parchite	1	16
59	Peña Enamorados	1	32
60	Periana	1	16
61	Pizarra	1	16
62	Rincón de la Victoria	1	16
63	Riogordo	1	32
64	Ronda CE	1	32
65	SP Alcántara	1	32
66	Tolox Millanas	1	16
67	Torrox	1	32
68	Vegueta Grama	1	16
69	Vélez Málaga	1	16
70	Villanueva Tapia	1	16
71	Viñuela	1	16
72	Yedra (La)	1	32

Siguiendo con el ejemplo de la estación Agujero y para el intervalo de escala obtenido, se procede a la obtención de la función exponente escaladora de momentos $K(q)$ (ecuación 2.5). La Figura 4.2 representa la función escaladora de momentos empíricos $K(q)$ donde, en el eje de abscisas, se representan los valores de los momentos q mientras que, en el eje de ordenadas, se representan los valores de dicha función.

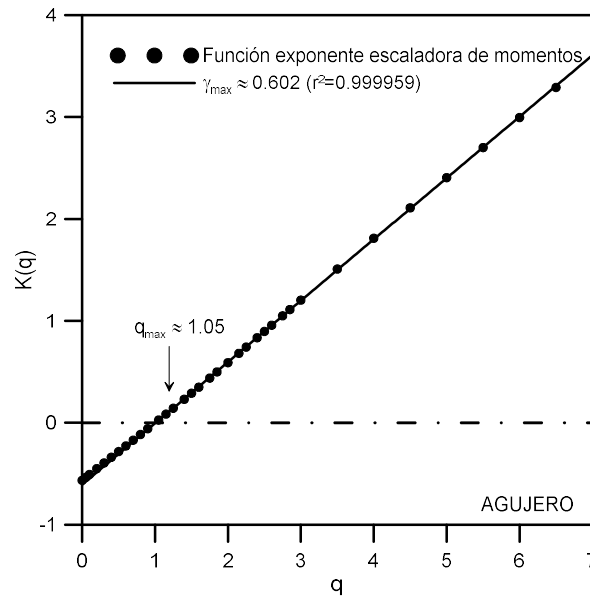


Figura 4.2. Función exponente escaladora de momentos. Estación Agujero

Se observa cómo la forma de la función $K(q)$ es prácticamente una línea recta, lo que indica un comportamiento monofractal de los datos de precipitación diaria para la estación considerada. El momento a partir del cual se observa un comportamiento estrictamente lineal viene marcado por $q > q_{crit}$ siendo q_{crit} el momento crítico (1,05 para el caso de la estación Agujero).

La Tabla 4.4 muestra los valores de q_{crit} obtenidos en cada una de las estaciones. Si se ordenan de manera ascendente se observa que el valor de q_{crit} oscila entre 0,0 y 3,5. El valor de q_{crit} que marca el límite entre la monofractalidad y la multifractalidad del proceso de lluvia es 2,0 (Tessier *et al.*, 1996).

Tabla 4.4. Valores de q_{max} para las distintas estaciones meteorológicas

ID	Estación	q_{crit}	ID	Estación	q_{crit}
1	Agujero	1,05	37	Conde Guadalhorce	2,50
2	Alcaucín Cortijo	2,40	38	Contaderas Forestal	2,85
3	Alcaucín Forestal	2,00	39	Corchado Central	1,25
4	Alfarnate	1,85	40	Cuevas Becerro	2,85
5	Alhaurín El Grande	2,15	41	El Burgo	3,00
6	Aljaima	2,15	42	Fuente Piedra	1,75
7	Almargen	2,85	43	Fuente Piedra Herriza	3,00
8	Almogía los Llanes	2,25	44	Gobantes Vivero	2,75
9	Alora	3,50	45	Humilladero	3,00
10	Alozaina	2,00	46	Hundidero	2,15
11	Alpandeire	1,85	47	Istán	2,85
12	Antequera el Águila	1,85	48	Jimena Líbar CE	1,85
13	Archidona	2,15	49	Las Mellizas	3,00
14	Arriate	2,75	50	Málaga Oficina	2,50
15	Benahavis	1,85	51	Marbella Inst. Laboral	1,00
16	Benalmádena	2,75	52	Mijas Faro Cal	1,15
17	Benamargosa	0,80	53	Moclinejo	1,85
18	Benamocarra	1,75	54	Montejaque	2,85
19	Benaoján CP	2,60	55	Nerja	1,40
20	Bobadilla Estación	1,85	56	Ojén	2,50
21	Borregos	1,25	57	Parauta Taramal	3,50
22	Buitreras CE	2,85	58	Parchite	2,60
23	Buitreras Presa	0,00	59	Peña Enamorados	0,80
24	Campillos	3,00	60	Periana	0,30
25	Canillas Aceituno	2,50	61	Pizarra	1,60

ID	Estación	q_{crit}	ID	Estación	q_{crit}
26	Cartajima	0,00	62	Rincón de la Victoria	3,50
27	Cartama Estación	1,50	63	Riogordo	3,00
28	Casabermeja	2,25	64	Ronda CE	2,85
29	Casabermeja VP	1,75	65	SP Alcántara	2,75
30	Casapalma	2,85	66	Tolox Millanas	0,40
31	Casarabonela	1,85	67	Torrox	1,60
32	Casares	1,85	68	Vegueta Grama	0,80
33	Chorro Estación	3,00	69	Vélez Málaga	3,50
34	Coin	1,50	70	Villanueva Tapia	3,00
35	Colmenar	2,25	71	Viñuela	3,00
36	Cómpeta	2,75	72	Yedra	3,50

Los valores obtenidos para el momento crítico (q_{crit}) servirán de base para delimitar las subregiones. Considerando el límite entre los procesos mono y multifractales, se han formado dos grupos o subregiones, resultando ambas heterogéneas en base al valor de H obtenido. A partir de estos resultados se han hecho diversas pruebas de agrupación en subregiones obteniéndose varias combinaciones hasta que los valores del estadístico H obtenidos han resultado satisfactorios (región homogénea o posiblemente heterogénea). Los resultados finales de todo este proceso se recogen en la Tabla 4.5.

Tabla 4.5. Identificación de regiones

Región	q_{crit}	Nº Estaciones	Estaciones	H
1	$1,05 < q_{crit} < 1,75$	12	Agujero, Benamocarra, Borregos, Cártama, Casabermeja, Venta Pineda, Coin, Corchado, Fuente Piedra, Mijas, Nerja, Pizarra y Tolox	0,06 (Homogénea)

Región	q_{crit}	Nº Estaciones	Estaciones	H
2	$2,6 < q_{crit} < 3,5$	27	Almargen, Alora, Arriate, Benalmádena, Benaolán CP, Buitreras CE, Campillos, Casapalma, Chorro, Cómpeta, Contaderas Forestal, Cuevas Becerro, El Burgo, Fuente Piedra Herriza, Gobantes Vivero, Humilladero, Istán, Las Mellizas, Montejaque, Parauta Taramal, Parchite, Riogordo, Ronda, SP Alcántara, Vélez Málaga, Villanueva Tapia y Viñuela	0,59 (Homogénea)
3	$0 < q_{crit} < 0,4$	4	Buitreras Presa, Cartajima, Periana y Tolox	1,21 (Posiblemente Heterogénea)
4	$0,8 < q_{crit} < 1,75$	16	Agujero, Benamargosa, Benamocarra, Borregos, Cártama, Casabermeja, Venta Pineda, Coin, Corchado Central, Fuente Piedra, Marbella, Mijas, Nerja, Peña Enamorados, Pizarra, Torrox y Vegueta Grama	1,28 (Posiblemente Heterogénea)

En base a los valores de q_{crit} , existe un cierto solape entre la Región 1 y la Región 4. Por ello a la hora de calcular los cuantiles de precipitaciones máximas en las estaciones comunes a ambas regiones, se considerarán los datos correspondientes a la región homogénea (Región 1). Para el resto de estaciones, se atenderá a los resultados de la Región 4. A pesar de la posible heterogeneidad de la región, Hosking y Wallis (1997) recomiendan el uso del análisis regional frente al local a la hora del cálculo de los diversos cuantiles.

Aquellas estaciones cuyos valores de q_{crit} oscilan 1,85 y 2,50 no forman región ni homogénea ni posiblemente heterogénea. Por lo tanto con todas ellas, junto con las estaciones Rincón de la Victoria y La Yedra (25 estaciones en total), se procede a la realización de grupos o regiones mediante análisis clúster.

Los parámetros elegidos para agrupar las estaciones son el valor de q_{crit} , la pendiente del tramo recto (γ_{max}) (ecuación 2.6) obtenido en la función $K(q)$ (ecuación 2.5) para valores de $q > q_{crit}$ y la precipitación máxima diaria para un periodo de retorno de 100 años (P_{100}). El valor de la precipitación máxima para un periodo de retorno de 100 años (P_{100}) se obtiene mediante aplicación del análisis local de frecuencias a las series de datos de lluvia máxima diaria anual en cada una de las 25 estaciones. La función de distribución elegida para el ajuste de las series de datos máximos es la Gumbel, función de dos parámetros y caso particular de la función de tres parámetros General de Valores Extremos, y el método empleado para formar los grupos es el K-means clustering. Los grupos formados aparecen en la Tabla 4.6 donde se recogen también los valores obtenidos para el estadístico H . La región 5, con un valor de H de 0,77 se considera homogénea, mientras que las estaciones de la región 6 no forman grupo homogéneo.

Tabla 4.6. Regiones formadas mediante Análisis Clúster

Región	Nº Estaciones	Estaciones	H
5	12	Casabermeja, Casarabonela, Casares, Colmenar, Conde de Guadalhorce, Hundidero, Jimena Líbar CE, Málaga Oficina, Moclinejo, Ojén, Rincón de la Victoria y Yedra.	0,77 (Homogénea)
6	13	Alcaucín Cortijo, Alcaucín Forestal, Alfarnate, Alhaurín El Grande, Aljaima, Almogia los Llanes, Alozaina, Alpandeire, Antequera el Águila, Archidona, Benehavis, Bobadilla Estación y Canillas Aceituno.	No forman Región

Una vez obtenidas las cinco subregiones homogéneas (Regiones 1, 2 y 5) o posiblemente heterogéneas (Regiones 3 y 4) en las que se puede dividir la provincia de

Málaga, el siguiente paso consiste en determinar la función o funciones de distribución a las que mejor se ajustan los datos de precipitaciones máximas de cada región.

A partir de los coeficientes lineales de variación, sesgo y curtosis, se calculan los valores del estadístico Z^{DIST} (ecuación 4.17) para cada región y mediante las funciones de distribución Logística Generalizada, General de Valores Extremos, Normal Generalizada, Pearson Tipo III y Pareto.

En la Tabla 4.7 aparecen marcadas mediante asterisco todas las funciones que cumplen con el criterio de bondad para el estadístico Z^{DIST} . Para cada región, la función más apropiada será aquella cuyo valor absoluto de Z^{DIST} sea más próximo a cero. Se observa que la función de distribución Pearson Tipo III es común para todas las subregiones, excepto para la subregión 2, en cuyo caso se aplicará la función de distribución General de Valores Extremos.

Tabla 4.7. Resultados obtenidos de las funciones de distribución (logística generalizada (A), general de valores extremos (B), normal generalizada (C), Pearson tipo III (D) y Pareto (E)

Regiones	A	B	C	D	E
1	4,75	1,65	1,50 *	0,74 *	-4,94
2	4,62	0,93 *	0,29 *	-1,21 *	-7,33
3	3,04	0,94 *	0,98 *	0,63 *	-3,38
4	5,24	1,74	1,48 *	0,52 *	-5,78
5	3,49	0,32 *	0,35 *	-0,21 *	-6,23

Una vez seleccionada la función de distribución más apropiada para cada región pueden obtenerse los cuantiles máximos diarios para distintos periodos de retorno en cada estación y a partir de las curvas regionales de frecuencia obtenidas para cada región con la función de distribución seleccionada. La Tabla 4.8 muestra los valores de las curvas regionales de frecuencia para las distintas subregiones.

Tabla 4.8. Curvas regionales de frecuencia para cada región

Región/Función de distribución	2	5	10	25	50	100
1/ Pearson Tipo III	0,94	1,29	1,51	1,77	1,95	2,13
2/ Gen Normal	0,94	1,25	1,46	1,74	1,94	2,14
3/ Pearson Tipo III	0,96	1,26	1,44	1,66	1,81	1,96
4/ Pearson Tipo III	0,94	1,29	1,51	1,78	1,97	2,15
5/ Pearson Tipo III	0,95	1,28	1,48	1,73	1,89	2,06

4.1.2.1. Conclusiones

La realización del Análisis Regional de Frecuencias de las precipitaciones máximas diarias anuales de la provincia de Málaga ha dado como resultado la heterogeneidad de la región compuesta por las 72 estaciones consideradas. A la vista de este resultado, se ha procedido a la división en subregiones. Para ello se han empleado diversos parámetros multifractales de cada una de las series de datos disponibles. Para la caracterización multifractal se ha analizado la función exponente escaladora de momentos, que proporciona información sobre la monofractalidad o multifractalidad de la lluvia. Dos parámetros procedentes de esta función, momento crítico y pendiente máxima, se han usado para la formación de subregiones. En total, se han formado cuatro subregiones, dos homogéneas y dos posiblemente heterogéneas.

Para agrupar en subregiones aquellas estaciones que no se han incluido en los grupos indicados anteriormente se ha efectuado un Análisis Clúster considerando los citados parámetros multifractales y el valor de la precipitación máxima diaria para un periodo de retorno de 100 años. Gracias a ello se han obtenido dos nuevos grupos, uno compuesto por 12 estaciones, homogéneo desde el punto de vista de las precipitaciones máximas anuales, y otro cuyas estaciones no forman región homogénea. Para el cálculo de los cuantiles, en este caso, se deberá recurrir al Análisis Local de Frecuencias.

En base a lo indicado se puede concluir que la caracterización multifractal de las series de datos de precipitación puede utilizarse en el Análisis Regional de Frecuencias para delimitar regiones homogéneas, bien directamente o como base para la agrupación de individuos mediante técnicas de clustering.

4.1.3. Formación de regiones homogéneas de precipitación máxima diaria anual en base a los valores de los parámetros γ_{max} y $K(0)$

Tras los resultados obtenidos en el apartado 4.1.2 se pretende ahora mejorar la agrupación de estaciones en base a una nueva combinación de parámetros multifractales característicos de las series de precipitación disponibles.

Para ello se trabaja en esta ocasión con las series de lluvia diaria disponibles para un total de setenta y cuatro estaciones meteorológicas distribuidas a lo largo de la región de Biobío de Chile. La información de estas se recoge en el apartado 3.1, la Figura 3.3 y las Tabla 3.3, y los resultados del control de calidad realizado, en el apartado 3.2.2.2.

Inicialmente y partiendo de las series de precipitaciones de las setenta y cuatro estaciones meteorológicas ubicadas en la Región de Biobío (Tabla 3.3), se han obtenido las series de precipitación máxima diaria anual. Cada serie de datos se ha caracterizado con los valores de los L-momentos, mostrados en la Tabla 4.9.

Tabla 4.9. Valores de los L-momentos para las estaciones analizadas en la región de Biobío

ID	Estación	L_1	$L_2 (\sigma)$	$L_3 (C_s)$	$L_4 (C_k)$	$L_5 (\lambda_5)$
1	Andalien	87,47	21,6359	0,2654	0,0930	-0,1060
2	Caman	131,89	19,2335	-0,1433	-0,0243	0,0152
3	Canal de la Luz	56,44	11,4600	0,5271	0,7382	0,7103
4	Cancha los Litres	80,27	21,8519	0,3663	0,2011	0,0859
5	Cañete	61,55	12,2202	0,2567	0,1413	0,0336
6	Caracol	139,98	22,0821	0,0415	0,0597	0,0554
7	Carampangue	55,59	5,0962	0,1960	0,2223	-0,0628
8	Cerro el Padre	104,96	15,0617	0,1195	0,1648	0,0515
9	Cerro Verde	67,97	8,7400	0,5980	0,7674	0,6407
10	Chillán Sendos	61,88	11,6006	0,3259	0,2479	0,0460
11	Chillán Viejo	79,95	16,0592	0,2773	0,1619	0,0702
12	Chillancito	83,17	20,1157	0,3411	0,2005	0,0927
13	Cholguán	81,76	17,8088	0,2605	0,1910	0,0606

ID	Estación	L ₁	L ₂ (σ)	L ₃ (C _s)	L ₄ (C _k)	L ₅ (λ_5)
14	Coelemu	64,74	14,1934	0,2880	0,2084	0,0670
15	Coihueco embalse	94,64	18,6473	0,2493	0,2497	0,1506
16	Concepción DGA	78,94	19,4902	0,2686	0,1135	0,0444
17	Concepción Edif MOP	94,31	19,2429	0,3565	0,0549	0,1317
18	Contulmo	75,97	15,9772	0,1889	0,2244	0,1319
19	Curanilahue	117,82	21,6507	0,1609	0,1603	0,0285
20	Dichato	65,35	13,1117	0,5015	0,3198	0,1568
21	Diguillín	120,95	18,4615	0,2310	0,1776	0,0961
22	Duqueco	59,08	12,5402	0,1350	0,3085	0,0801
23	Embalse Pangue	160,09	29,9952	0,4236	0,1114	-0,3061
24	Embalse Ralco	99,57	14,7267	-0,1347	0,1727	-0,6892
25	Estero Bellavista Tome	55,66	4,9900	-0,0902	0,2685	-0,7495
26	Estero Hualqui Desembocadura	-	-	-	-	-
27	Estero Nonguen	62,86	10,5700	0,5024	0,0114	-0,3254
28	Estero Plegarias	92,90	12,8900	-0,0101	0,0962	1,1792
29	Estero Quilque	48,03	7,2267	0,2565	0,1513	-0,0923
30	Fundo Atacalco	124,99	19,4215	0,0779	0,2228	0,0337
31	Fundo Bella Rosa	65,50	9,5333	0,0000	-0,5734	0,0000
32	General Cruz	49,94	4,4867	0,4506	0,3380	0,2931
33	La Punilla	111,81	19,4005	0,0842	0,1368	0,0644
34	Laja	61,64	11,2773	0,3002	0,3597	0,0948
35	Las Achiras	67,62	12,6771	0,2048	0,1389	0,0664
36	Las Cruces	121,72	20,3219	0,0872	0,0510	-0,0399
37	Las Pataguas	75,98	18,5976	0,4535	0,1423	-0,0232
38	Las Trancas	120,32	22,2379	0,1184	0,1204	-0,0003
39	Los Ángeles	73,76	14,4605	0,2902	0,1844	0,0739
40	Mangarral	72,72	17,4929	0,3688	0,2635	0,1323
41	Mayulermu	100,21	17,8600	0,1599	0,1826	0,0387
42	Millauquen	72,16	16,0314	0,3042	0,1427	0,0717
43	Mulchen	79,95	13,4814	0,2954	0,2225	0,0510
44	Nueva Aldea	70,35	16,1904	0,2781	0,1772	0,0398

ID	Estación	L_1	$L_2 (\sigma)$	$L_3 (C_s)$	$L_4 (C_k)$	$L_5 (\lambda_5)$
45	Parque Nahuelbuta	80,13	17,3598	0,1345	0,1188	0,0169
46	Pemuco	73,10	13,5983	0,1272	0,1412	0,0417
47	Pilguen	88,26	15,5524	0,2773	0,1429	0,0461
48	Portezuelo	-	-	-	-	-
49	Quilaco	81,98	12,6511	0,1208	0,1999	0,0550
50	Quillaileo	103,65	15,3175	-0,0897	0,0171	0,0572
51	Rafael	86,57	21,8929	0,3377	0,2085	0,1830
52	Rio Bio Bio Coihue	-	-	-	-	-
53	Rio Bio Bio Desembocadura	60,18	10,1923	0,0663	0,1383	0,1103
54	Rio Bio Bio Esperanza	97,32	15,2900	0,1485	0,6305	-0,1491
55	Rio Bio Bio Llauquen	92,59	18,7636	0,0349	0,0507	0,0699
56	Rio Bio Bio Longitudinal	48,38	13,4100	0,7569	0,7427	0,9605
57	Rio Bio Bio Rucalhue	88,88	16,9551	0,2176	0,0630	0,0568
58	Rio Curanilahue Curanilahue	-	-	-	-	-
59	Rio Diguillín San Lorenzo Atacalco	95,64	15,1000	-0,1232	0,5497	0,0623
60	Rio Duqueco Cerrillos	64,88	15,2100	-0,0256	0,1026	-0,2051
61	Rio Itata Cuelemu	48,86	12,7500	0,6471	0,7686	0,6086
62	Rio Laja Salto	-	-	-	-	-
63	Rio Laja Tucapel	87,71	18,8449	0,3612	0,0609	-0,1098
64	Rio Leiva Puentecontulmo	62,28	5,4100	-0,2606	-0,3216	0,9945
65	Rio Ñuble San Fabián	95,19	19,1115	-0,1189	0,0509	-0,1073
66	San Agustín Puñual	74,19	16,3033	0,2282	0,1932	0,0841
67	San Carlos Puren	75,17	16,0739	0,2878	0,2458	0,1483
68	San Fabián	118,71	26,2438	0,1381	0,0929	0,0053
69	San Jose Munilque	66,15	11,0548	0,3020	0,1943	0,1295
70	San Lorenzo	125,54	22,6751	0,1007	0,2122	0,0360
71	San Lorenzo Bio Bio	104,23	14,9113	0,0884	0,1145	0,1421
72	Tirua	86,71	8,9038	0,0359	0,1661	0,2045
73	Trupán	87,58	14,1291	0,1946	0,0848	0,0225
74	Tucapel	87,75	11,9978	0,2825	0,2212	0,019

Tal como se muestra en la Tabla 4.9, las estaciones Estero Hualqui Desembocadura, Portezuelo, Río Biobío Coihue, Río Curanilahue y Río Laja Salto no han dado resultados, probablemente debido al número de datos tan escaso, por tanto, estas estaciones no se tendrán en cuenta y se considerarán un total de sesenta y nueve estaciones para llevar a cabo los procesos posteriores.

Se considera por lo tanto una primera región llamada Biobío, formada por sesenta y nueve estaciones meteorológicas, que será estudiada con el fin de analizar su homogeneidad. Siguiendo la metodología propuesta por Hosking y Wallis (1997), en regiones con más de 15 estaciones meteorológicas el valor crítico de discordancia D_i es 3,00 (Tabla 4.1).

Considerando este valor límite de discordancia, se encontraron siete estaciones discordantes en la primera prueba: Cerro Verde ($D_i=4,69$), Fundo Bella Rosa ($D_i=5,27$), General Cruz ($D_i=3,91$), Río Biobío Longitudinal ($D_i=3,65$), Río Diguillín ($D_i=4,59$), Río Itata ($D_i=3,21$) y Río Leiva Puentecontulmo ($D_i=3,26$); tres estaciones discordantes en la segunda prueba: Canal de la Luz ($D_i=7,42$), Estero Nonguen ($D_i=4,16$) y Río Biobío Esperanza ($D_i=5,72$); una estación discordante en la tercera prueba: Estero Bellavista ($D_i=3,69$); y una estación discordante en la cuarta prueba: Carampangue ($D_i=3,07$). Después de eliminar todas las estaciones discordantes del análisis, el valor del estadístico H fue de 1,79 para la región formada por cincuenta y siete estaciones (tras descartar las discordantes). Este valor indica que la región es posiblemente heterogénea, por tanto, deben formarse subregiones potencialmente homogéneas. Así pues, se ha llevado a cabo la caracterización multifractal de los datos de lluvia para todas las estaciones meteorológicas (69) mediante el escalado de momentos q . Como ejemplo, en la Figura 4.3 se muestra la representación doblemente logarítmica de los q -ésimos momentos de intensidad de lluvia para una de las estaciones de la Región de Biobío, Andalien en este caso.

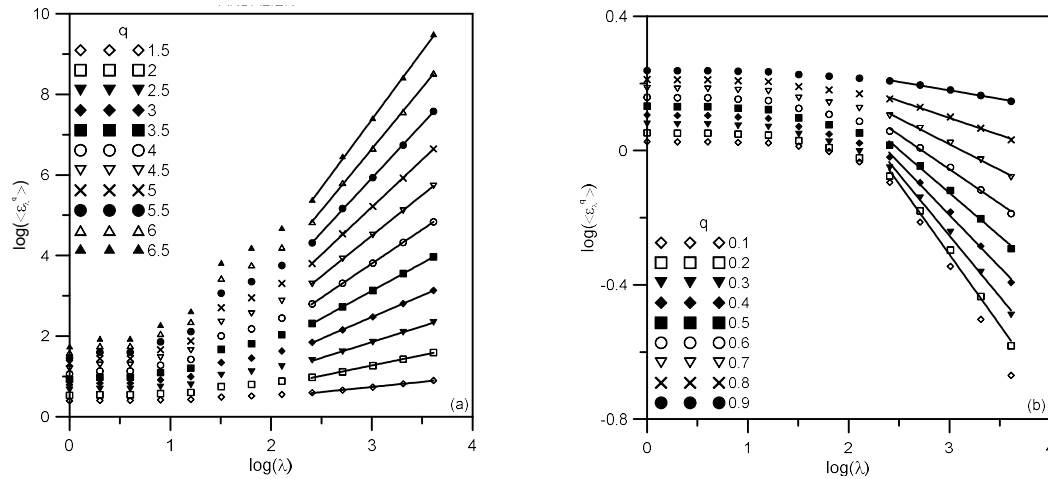


Figura 4.3. Representación log-log de los momentos q-ésimos. Estación Andalien.

La representación de los momentos de orden superior a uno se recoge en la parte izquierda de la Figura (a), mientras que la parte derecha (b) muestra los resultados para los momentos de orden inferior a la unidad. Se representan sólo algunos momentos q con el fin de aclarar la figura y los resultados. En todos los casos se ha detectado un comportamiento lineal (líneas continuas para los ajustes en ambas figuras) pero con límites diferentes. Para el resto de estaciones se ha detectado un comportamiento de escala a partir del cual se puede obtener la función exponente escaladora de momentos $K(q)$.

De esta forma, continuando con el análisis multifractal y teniendo en cuenta la ecuación (2.5), se ha obtenido la función $K(q)$ para las sesenta y nueve estaciones. La Figura 4.4 muestra la representación de la función exponente escaladora de momentos $K(q)$ para la estación de Andalien.

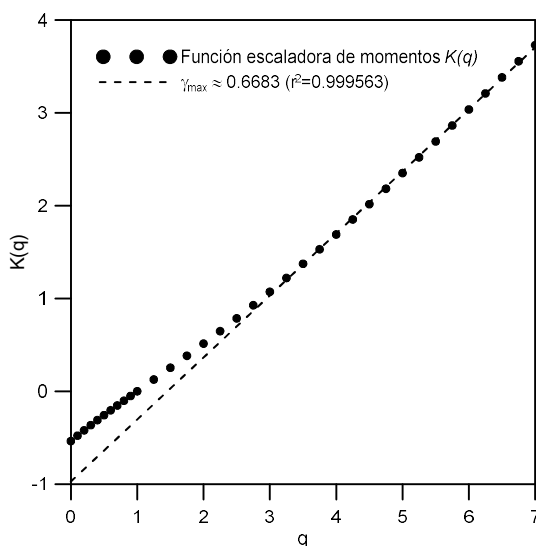


Figura 4.4. Función exponente escaladora de momentos. Estación Andalien.

En esta ocasión, la información que se obtendrá de la función $K(q)$ serán los valores de γ_{\max} y $K(0)$. El primero de ellos puede relacionarse con la presencia de valores extremos o raros en la serie analizada, mientras que el segundo proporciona información de la presencia de valores bajos o nulos en la serie.

La Tabla 4.10 muestra el valor de estos parámetros para cada una de las estaciones de la región.

Tabla 4.10. Parámetros multifractales obtenidos.

Estación	γ_{\max}	$K(0)$
Andalien	0,6683	-0,5370
Caman	0,5025	-0,5409
Canal de la Luz	0,6857	-0,3120
Cancha los Litres	0,7916	-0,4636
Cañete	0,6927	-0,3103
Caracol	0,4559	-0,0150
Carampangue	0,4587	-0,4985
Cerro el Padre	0,5596	-0,4572

Estación	γ_{\max}	$K(\theta)$
Cerro Verde	0,4429	-0,4751
Chillán Sendos	0,5810	-0,4160
Chillán Viejo	0,9199	-0,4629
Chillancito	0,8239	-0,5047
Cholguán	0,5602	-0,3281
Coelemu	0,6936	-0,5245
Coihueco embalse	0,6583	-0,4597
Concepción DGA	0,8480	-0,5000
Concepción Edif MOP	0,2304	-0,1200
Contulmo	0,6229	-0,4204
Curanilahue	0,6131	-0,3612
Dichato	0,8493	-0,4862
Diguillín	0,6262	-0,4499
Duqueco	0,7120	-0,5153
Embalse Pangué	0,5803	-0,0762
Embalse Ralco	0,5351	-0,4560
Estero Bellavista Tome	0,9041	-0,5210
Estero Nonguen	0,2826	-0,0279
Estero Plegarias	0,6634	-0,4051
Estero Quilque	0,6747	-0,5424
Fundo Atacalco	0,2647	0,0000
Fundo Bella Rosa	0,7184	-0,4053
General Cruz	0,5042	-0,4547
La Punilla	0,3849	-0,5014
Laja	0,6194	-0,4318
Las Achiras	0,6281	-0,4503
Las Cruces	0,6457	-0,5038
Las Pataguas	0,6565	-0,4702
Las Trancas	0,5239	-0,2144
Los Ángeles	0,7155	-0,4726
Mangarral	0,5246	-0,4056

Estación	γ_{\max}	$K(\theta)$
Mayulermu	0,6405	-0,4290
Millauquen	0,6475	-0,3890
Mulchen	0,7022	-0,4498
Nueva Aldea	0,4706	-0,1858
Parque Nahuelbuta	0,5492	-0,4596
Pemuco	0,3893	-0,3622
Pilguen	0,5303	-0,4368
Quilaco	0,5264	-0,3849
Quillaileo	0,4941	-0,2857
Rafael	0,8372	-0,5572
Rio Bio Bio Desembocadura	0,6632	-0,5019
Rio Bio Bio Esperanza	0,6010	-0,4717
Rio Bio Bio Llauquen	0,5116	-0,4611
Rio Bio Bio Longitudinal	0,7991	-0,4334
Rio Bio Bio Rucalhue	0,6945	-0,2620
Rio Diguillín San Lorenzo Atacalco	0,4181	-0,5250
Rio Duqueco Cerrillos	0,8123	-0,5172
Rio Itata Cuelemu	0,9125	-0,5983
Rio Laja Tucapel	0,8016	-0,5376
Rio Leiva Puentecontulmo	0,5469	-0,4121
Rio Ñuble San Fabián	0,6652	-0,5783
San Agustín Puñual	0,5875	-0,5300
San Carlos Puren	0,7141	-0,4563
San Fabián	0,4840	-0,4017
San Jose Munilque	0,4324	-0,1818
San Lorenzo	0,4454	-0,1276
San Lorenzo Bio Bio	0,5987	-0,4608
Tirua	0,2348	-0,4082
Trupán	0,4836	-0,4687
Tucapel	0,3577	-0,0905

El rango de valores de γ_{max} se encuentra entre el valor mínimo 0,2304 para la estación Concepción Edif MOP y el valor máximo 0,9199 para la estación Chillán Sendos. Estos valores pueden proporcionar información sobre los eventos raros y extremos de la serie temporal (Veneziano y Furcolo, 2002; García-Marín *et al.*, 2013b), tal y como se ha comentado anteriormente.

Los valores de $K(0)$ obtenidos (Tabla 4.10) varían entre -0,5983 para la estación Río Ítata Cuelemu y 0 en la estación Fundo Atacalco.

Con todos los valores de γ_{max} y $K(0)$ se procede a la realización de un análisis clúster (apartado 4.1.1.7) caracterizando por un vector de dos componentes $(\gamma_{max}, K(0))$ cada una de las estaciones analizadas.

En la primera prueba de análisis clúster se han formado dos grupos, uno de 26 estaciones y otro de 43. Para cada grupo obtenido, se calculan los distintos estadísticos del ARF. De acuerdo al valor del estadístico H obtenido para el primer grupo, $H=3,28$, debe descartarse. En las tres pruebas realizadas al segundo grupo para analizar su homogeneidad, aparecen dos estaciones discordantes en la prueba uno Fundo Bella Rosa ($D_i= 5,43$) y Río Biobío Longitudinal ($D_i= 3,04$); tres estaciones discordantes en la prueba dos: Canal de la Luz ($D_i= 3,64$), Río Biobío Esperanza ($D_i= 3,71$) y Río Ítata Cuelemu ($D_i=4,57$); y una estación discordante en la prueba tres: Estero Bellavista ($D_i=3,58$). Finalmente, tras eliminar las estaciones discordantes del análisis, se obtiene una primera región homogénea formada por 37 estaciones, cuyo valor del estadístico H es de 0,40 (Tabla 4.11).

Tabla 4.11. Región Homogénea 1

Región	Estaciones (37)	H
1	Andalien, Caman, Cancha los Litres, Cañete, Cerro El Padre, Chillán Sendos, Chillán Viejo, Chillancito, Coelemu, Coihueco Embalse, Concepción DGA, Contulmo, Curanilahue, Dichato, Diguillín, Duqueco, Embalse Ralco, Estero Plegarias, Estero Quilque, Laja, Las Achiras, Las Cruces, Las Pataguas, Los Ángeles, Mayulermu, Millauquen, Mulchen, Parque Nahuelbuta, Rafael, Río Biobío Desembocadura, Río Biobío Rucalhue, Río Duqueco Cerrillos, Río Laja Tucapel, Río Ñuble San Fabián, San Agustín Puñual, San Carlos Puren y San Lorenzo Biobío.	0,40

En la segunda prueba de análisis clúster se parte de las 32 estaciones restantes, descartadas de los grupos anteriores. Inicialmente se obtienen dos grupos de estaciones, uno de 10 y otro de 22 estaciones. Tras comprobar que ninguno de ellos resulta homogéneo (H grupo 1= 1,06 y H grupo 2= 1,92) se decide repetir el análisis clúster formando tres grupos. Los tres grupos obtenidos están formados por las siguientes estaciones: el primer grupo con 10 estaciones, el segundo grupo con 5 estaciones y el tercer grupo con 17 estaciones. Tras analizar los valores del estadístico H de cada uno de los grupos (H grupo 1= 1,06; H grupo 2= 0,02 y H grupo 3= 2,64) se obtiene la segunda región homogénea, formada por cinco estaciones (Tabla 4.12).

Tabla 4.12. Región Homogénea 2

Región	Estaciones (5)	H
2	Canal de la Luz, Estero Bellavista Tome, Fundo Bella Rosa, Río Biobío Longitudinal y Río Ítata Cuelemu	0,02

En la tercera prueba del análisis clúster se parte de las 27 estaciones restantes, descartadas de los grupos anteriores. Inicialmente se obtienen dos grupos de estaciones, uno de 10 y otro de 17 estaciones pero, tras comprobar que ninguno de ellos resulta homogéneo (H grupo 1= 1,06 y H grupo 2= 2,64) se decide repetir el mismo análisis clúster pero formando, en este caso, tres grupos. Los tres grupos obtenidos están formados por las siguientes estaciones: el grupo uno, 10 estaciones, el grupo dos, 11 estaciones y el grupo tres, 6 estaciones. Tras analizar los valores del estadístico H de cada uno de los grupos (H grupo 1= 1,06; H grupo 2= 2,87 y H grupo 3= 0,90) se obtiene que el grupo 3 forma la tercera región homogénea, formada por seis estaciones.

Tabla 4.13. Región Homogénea 3.

REGIÓN	ESTACIONES (5)	H
3	Carampangue, Cerro Verde, La Punilla, Pemuco, Río Diguillín y Tirua	0,90

En la cuarta prueba de análisis clúster se parte de las 21 estaciones restantes y se decide formar, directamente, tres grupos, obteniendo los siguientes resultados: un primer grupo formado por seis estaciones, un segundo grupo formado por 11 y un tercer grupo formado por 4. En este caso, el último grupo no se tiene en cuenta por tener un número de estaciones tan escaso. En este trabajo, se considera como región mínima aquella que dispone, al menos, de cinco estaciones meteorológicas.

Tras analizar el valor del estadístico H en los grupos uno y dos se obtienen los siguientes resultados: H grupo 1 = 0,73 y H grupo 2 = 2,87, obteniendo una cuarta región homogénea formada por seis estaciones (Tabla 4.14).

Tabla 4.14. Región Homogénea 4.

Región	Estaciones (5)	H
4	Caracol, Embalse Pangué, Las Trancas, Nueva Aldea, San José Munilque y San Lorenzo	0,73

En la quinta prueba de Análisis Clúster se parte de las 15 estaciones restantes y se decide formar cuatro grupos, resultando todos tan pequeños (grupo uno formado por cuatro estaciones, grupo dos formado por dos estaciones y grupo cuatro formado por cuatro estaciones) que ninguno de ellos resulta de útil. Tan sólo el grupo tres está formado por cinco estaciones. Tras analizar el valor del estadístico H en este tercer grupo se obtiene como resultado $H= 0,88$, resultando la última región homogénea, formada por cinco estaciones (Tabla 4.15).

Tabla 4.15. Región Homogénea 5.

Región	Estaciones (5)	H
5	Mangarral, Pilguen, Quilaco, Río Biobío Esperanza y Río Leiva Puentecontulmo	0,88

Tras analizar diversas pruebas con las diez estaciones restantes: Cholguán, Concepción Edif MOP, Estero Nonguen, Fundo Atacalco, General Cruz, Quillailleo, Río Biobío

Llauquen, San Fabián, Trupán y Tucapel, se llega a la conclusión de que no forman ninguna región homogénea. Para estas diez estaciones, el Análisis Regional de Frecuencias no proporciona buenos resultados, por lo que el Análisis Local de Frecuencias será la opción para obtener los cuantiles de lluvia cuando sea necesario.

4.1.3.1. Conclusiones

Esta parte de la Tesis presenta un avance a la metodología para la agrupación de estaciones cuando se realiza un Análisis Regional de Frecuencias con datos de precipitación máxima diaria anual que se mostró en el apartado 4.1.2.

En base a los resultados obtenidos, se demuestra que la caracterización multifractal de los datos de lluvia validados resulta un método muy útil para agrupar series de datos de precipitación diaria en regiones homogéneas cuando se realiza un Análisis Regional de Frecuencias con este tipo de datos meteorológicos. En este caso, usando simplemente valores característicos del comportamiento multifractal de la lluvia. La novedad de este estudio radica en el uso de nuevos parámetros multifractales de la lluvia diaria para formar las nuevas regiones. El resto de las características de uso común en metodologías de análisis regional (latitud, longitud, índice de lluvia...) no fueron necesarias para la agrupación de las estaciones.

El método propuesto por Hosking y Wallis (1997) se ha aplicado, en primer lugar, a la totalidad de las estaciones meteorológicas disponibles para la zona de estudio (Región Biobío de Chile). Los resultados no han sido satisfactorios y la región se ha dividido en sub-regiones. Cada estación ha sido caracterizada en este caso por dos valores obtenidos a partir de la función exponente escaladora de momentos $K(q)$, γ_{\max} y $K(0)$. El primero de ellos está relacionado con la presencia de precipitaciones extremas o eventos raros en el conjunto de datos. El segundo está relacionado con valores bajos o nulos de lluvia. La diferencia en estos valores ha sido la base para el proceso de agrupación mediante Análisis Clúster. Se han obtenido así cinco regiones homogéneas y un último grupo heterogéneo formado por diez estaciones. Para estas últimas estaciones, el análisis local de frecuencias será la mejor opción para el cálculo de cuantiles de precipitación. Para el resto de las estaciones de la Región de Biobío (Chile), el ARF podrá utilizarse sin problema para obtener datos de precipitación máxima diaria correspondientes a cualquier periodo de retorno.

4.1.4 Obtención de regiones homogéneas de precipitación a partir de los valores de la dimensión fractal generalizada

En esta parte de la tesis el objetivo es la formación de regiones homogéneas de precipitación máxima diaria anual, pero esta vez utilizando para la agrupación los resultados de los valores de la dimensión fractal de los datos de lluvia diaria disponibles en las 53 estaciones de la región del Maule (Chile) (Véase apartado 3.1).

4.1.4.1. Análisis multifractal: aplicación del formalismo del StrangeAtractor

Una vez validados los datos iniciales (véase apartado 3.2.2.2), se procede a analizar el comportamiento multifractal de los mismos a partir del valor de su dimensión fractal. Para ello se aplica el formalismo del Strange Atractor, descrito en detalle en el apartado (2.2.2), a cada una de las series validadas obteniéndose el espectro de Renji (ecuación 2.15) para cada localización. La Figura 4.5 muestra las dimensiones fractales generalizadas D_q para valores de q que oscilan entre -10 a 10, para 4 localidades: Agua Fría, Bullileo, Liguay y San Rafael. Como puede observarse, los valores de la dimensión fractal disminuyen al aumentar los valores de q , lo que confirma la naturaleza multifractal de las series analizadas.

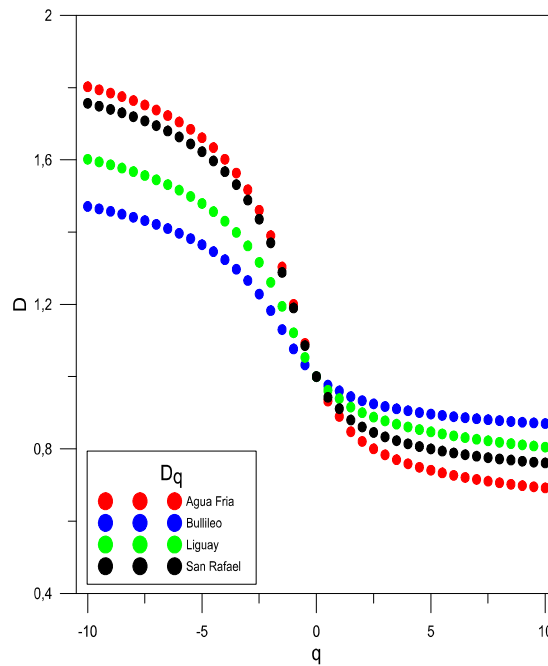


Figura 4.5. Valores de la dimensión fractal generalizada para 4 estaciones de la región del Maule

Los valores de D_0 son 1 para todas las series de datos analizadas, lo que pone de manifiesto que se está trabajando en una dimensión. Para valores inferiores y mayores de q , los valores de q obtenidos son diferentes a la unidad, cumpliéndose en todos los casos $D_0 > D_1 > D_2$. Para valores de q inferiores a 0 los valores más elevados de D_q se han obtenido para la estación de Agua Fría, seguida por San Rafael y Liguay, y estando los valores más bajos en la estación de Bullileo. Para valores de q superiores a 0, el comportamiento es el opuesto, siendo los valores obtenidos para Bullileo los mayores, seguidos por Liguay y San Rafael, y contando esta vez Agua Fría con los valores de D_q más bajos.

La Tabla 4.16 recoge los valores de D_1 y D_2 para todas las series de datos analizadas. D_1 proporciona una medida del grado de heterogeneidad (Davis *et al.*, 1994) y caracteriza la intensidad y la distribución de las singularidades presentes en los datos respecto a la media (Ariza-Villaverde *et al.*, 2013).

Tabla 4.16. Valores de las dimensiones fractales D_1 y D_2 , así como del grado de multifractalidad (MD) para las series de datos analizadas (ID) en la región del Maule

ID	D_1	D_2	MD	ID	D_1	D_2	MD
1	0,889615	0,820564	0,920359	28	0,937466	0,893687	0,584380
2	0,955456	0,925041	0,507552	29	0,974909	0,955247	0,299732
3	0,942059	0,905053	0,725544	30	0,955529	0,925813	0,477947
4	0,960375	0,933376	0,468767	31	0,956797	0,923613	0,413641
5	0,974106	0,954068	0,272129	32	0,953460	0,914992	0,377158
6	0,980991	0,965033	0,204177	33	0,951121	0,911541	0,416012
7	0,879698	0,802271	1,224827	34	0,978045	0,959457	0,236557
8	0,959815	0,926981	0,336645	35	0,883600	0,822819	1,481674
9	0,958850	0,930099	0,392157	36	0,919515	0,872760	0,994822
10	0,919259	0,864137	0,574581	37	0,913781	0,857483	0,899522
11	0,951433	0,913196	0,451888	38	0,918115	0,860271	0,833337
12	0,973490	0,953580	0,321003	39	0,892489	0,823538	1,143640
13	0,920847	0,873019	0,834535	40	0,909142	0,859638	0,951834
14	0,950020	0,901733	0,390810	41	0,945099	0,901751	0,500260
15	0,954170	0,919747	0,398820	42	0,877382	0,793721	1,108003
16	0,982265	0,966985	0,197977	43	0,918266	0,867041	0,823724
17	0,972172	0,948569	0,263800	44	0,955153	0,918287	0,437054
18	0,970621	0,945680	0,359593	45	0,961036	0,931249	0,376100
19	0,966911	0,942063	0,345346	46	0,980835	0,964870	0,199753
20	0,978528	0,962984	0,290351	47	0,911176	0,860429	0,822637
21	0,916204	0,864462	0,862946	48	0,977476	0,960844	0,279331
22	0,984788	0,969899	0,137865	49	0,958848	0,925689	0,430586
23	0,938829	0,900063	0,632138	50	0,929421	0,883229	0,554809
24	0,938470	0,901135	0,608994	51	0,930459	0,890896	0,767668
25	0,921801	0,874771	0,864936	52	0,899884	0,833084	0,759191
26	0,966235	0,941503	0,396674	53	0,899037	0,840771	0,958326
27	0,972364	0,949774	0,317954				

El valor más bajo de D_I es el que se obtiene para la estación Río Maule Salto (0,877382) lo que pone de manifiesto que las singularidades de la serie de datos se distribuyen de forma más dispersa que en el resto de las series analizadas para las que se obtienen valores mayores de D_I y por lo tanto puede esperarse una distribución más homogéneas de las singularidades. El mayor valor de D_I se obtiene para la estación La Sexta (0,984788).

Los valores de la dimensión de correlación (D_2) se muestran también en la Tabla 4.16, donde los valores más bajos y más altos se han obtenido para las mismas estaciones que D_I , siendo 0,793721 y 0,969889 para Río Maule Salto y La Sexta, respectivamente. La dimensión de correlación describe la probabilidad de encontrar datos que, dentro del conjunto de datos analizado, se encuentren a una determinada ‘distancia’ de un dato que también pertenece al conjunto analizado (Ariza-Villaverde *et al.*, 2013).

Una vez analizado el comportamiento multiescalar de las series de datos de lluvia a partir de la función D_q , se obtienen los espectros multifractales $f(\alpha)$ (ecuación 2.18) para todas las series. La Figura 4.6 muestra los espectros multifractales para las mismas estaciones de la Figura 4.5. Para todas ellas, el espectro tiene forma de parábola invertida. El espectro permite cuantificar con más detalle las propiedades de correlación y proporciona información de varios exponentes fractales presentes en las series analizadas. Su anchura es una medida del grado de multifractalidad (MD , del inglés ‘Multifractal Degree’) de la serie analizada (Telesca y Lovullo, 2011). A mayor anchura del espectro, mayor será la multifractalidad del proceso analizado. Para un conjunto de datos monofractal, la anchura del espectro será cero (Maity *et al.*, 2015). Según Telesca *et al.* (2004) la anchura del espectro puede obtenerse a partir de la función D_q , como la diferencia entre los valores D_{-5} y D_5 . A mayor grado de multifractalidad (MD) mayor heterogeneidad de los datos analizados. La Tabla 4.16 recoge los valores del grado de multifractalidad (MD) para todas las estaciones analizadas, siendo los valores mínimos y máximos de 0,137865 para la estación La Sexta y 1,481674 para la estación Río Ancoa, respectivamente. Para las estaciones que aparecen en la Figura 4.6, el mayor valor del grado de multifractalidad se obtiene para Agua Fria (0,920359), seguida por San Rafael (0,822637) y Liguay (0,632138). El valor más bajo del grado de multifractalidad corresponde a la estación de Bullileo (0,468812).

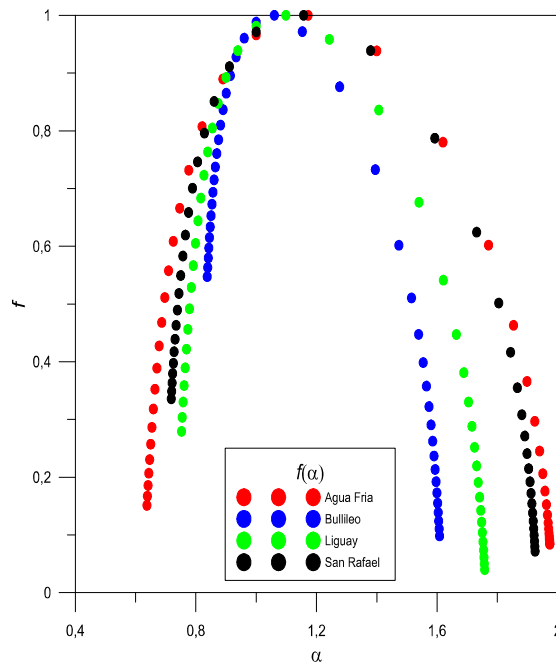


Figura 4.6. Espectro multifractal $f(\alpha)$ para las estaciones de Agua Fría, Bullileo, Liguay y San Rafael.

A partir de la forma del espectro multifractal (Figura 4.6) puede obtenerse también cierta información de los datos analizados. Los espectros más anchos y redondeados corresponden a una distribución de valores con una mayor variabilidad. Los espectros obtenidos para Agua Fría y Bullileo tienen formas diferentes, siendo el obtenido para Agua Fría más redondeado y más ancho que el obtenido para Bullileo. Este comportamiento diferente en ambas series de datos se había detectado anteriormente por García-Marín *et al.* (2015) estando relacionado con el porcentaje de días sin lluvia así como con la presencia de eventos raros y extremos en las series temporales analizadas.

4.1.4.2. Formación de regiones homogéneas

Una vez analizado el comportamiento multifractal de las series de datos diarios de precipitación en la región del Maule, el siguiente paso consiste en el estudio de la posibilidad de formar regiones homogéneas de precipitación máxima diaria anual en base a dicho comportamiento. Para ello, a partir de las series de datos de lluvia diaria previamente validadas (véase apartado 3.2.2.2) se han obtenido las series de

precipitaciones máximas diarias anuales en cada una de las cincuenta y tres estaciones disponibles. A continuación, cada una de las estaciones se ha caracterizado por el valor de los L-Momentos de las series de datos máximos anuales y sus cocientes ($L-C_v$, $L-C_s$ y $L-C_k$) (Tabla 4.17).

Tabla 4.17. L-Momentos para las 53 estaciones analizadas

ID	$L-C_v$	$L-C_s$	$L-C_k$
1	0,17021	0,02470	0,02980
2	0,19991	0,16800	0,18040
3	0,17742	0,03650	0,07030
4	0,16041	0,06240	0,10440
5	0,20386	0,24130	0,22860
6	0,15524	0,09070	0,20070
7	0,16360	0,02740	0,11080
8	0,18376	0,10180	0,14270
9	0,16922	0,16470	0,15940
10	0,23922	0,24720	0,11410
11	0,13473	0,09070	0,05050
12	0,17035	0,21500	0,15470
13	0,18315	0,15410	0,20060
14	0,22194	0,17490	0,06870
15	0,19496	0,12300	-0,00780
16	0,21462	0,21770	0,16540
17	0,22309	0,16160	0,12500
18	0,17780	0,24470	0,14250
19	0,14048	0,03500	0,23520
20	0,16596	0,15820	0,18680
21	0,12567	0,02960	0,11540
22	0,22131	0,27010	0,10090
23	0,16202	0,38420	0,33930
24	0,15056	0,11300	0,08560
25	0,19731	0,21080	0,21990
26	0,23472	0,25290	0,28920
27	0,21118	0,23650	0,17940
28	0,17386	0,06220	0,09350
29	0,22222	0,24950	0,12030
30	0,19221	0,18550	0,11790
31	0,14389	0,14170	0,16590

ID	$L-C_v$	$L-C_s$	$L-C_k$
32	0,15280	0,05750	0,11810
33	0,19044	0,23040	0,11650
34	0,18842	0,36920	0,42320
35	0,18862	-0,00410	0,23530
36	0,13467	0,09730	0,28580
37	0,12764	-0,08870	0,31450
38	0,20819	0,16120	0,00810
39	0,15165	0,01650	-0,03490
40	0,21089	0,29450	0,20090
41	0,19334	-0,01480	0,01010
42	0,19528	0,02530	-0,00380
43	0,18205	0,01940	-0,07320
44	0,20574	0,07030	0,05610
45	0,16841	0,11990	0,13870
46	0,16179	0,08170	0,06450
47	0,07668	0,07870	0,08400
48	0,18966	-0,01380	0,06250
49	0,14436	0,20050	0,32820
50	0,15721	0,07360	0,09140
51	0,16742	0,10020	0,12600
52	0,18563	0,15510	0,28900
53	0,23550	0,27310	0,28130

Con todos los valores de los L-momentos (Tabla 4.17) se ha realizado un ARF de las precipitaciones máximas diarias anuales para una primera gran región llamada Maule. Siguiendo el criterio de Hosking y Wallis (1997) todas las estaciones para las que el valor de la Discordancia (D_i) sea mayor que 3,00 (Tabla 4.1), han de eliminarse de la región. Cinco de las estaciones inicialmente incluidas en la región Maule cumplen este criterio y han de ser eliminadas del análisis inicial: Agua Fria ($D_i = 3,35$), Quella ($D_i = 3,01$), Río Ancoa ($D_i = 3,96$), Río Loncomilla ($D_i = 5,04$) y San Rafael ($D_i = 4,66$). Para la región Maule, ahora compuesta por cuarenta y ocho estaciones, se calcula a continuación el valor del estadístico H , siendo de 2,35 (Tabla 4.18). Considerando que el valor obtenido para la región Maule es superior a 2, la región se considera heterogénea y se plantea la necesidad de descomponerla en subregiones que puedan ser

potencialmente homogéneas desde el punto de vista de la precipitación máxima diaria anual.

Tabla 4.18. Resultados de los tests de homogeneidad para las distintas regiones formadas

Región	Estaciones	Discordantes	H
Maule	Todas	23,34,35,37,47	2,35
Region 1	1, 7, 10, 13, 21, 25, 35, 36, 37, 38, 39, 40, 42, 43, 47, 50, 51, 52, 53.	47	0,33
Region 2	2, 3, 4, 8, 9, 11, 14, 15, 23, 24, 28, 30, 31, 32, 33, 41, 44, 45, 47, 49	23, 47, 49	0,76
Region3	5, 6, 12, 16, 17, 18, 19, 20, 22, 23, 26, 27, 29, 34, 46, 47, 48, 49	47	0,80

La división en subregiones se apoya en los valores obtenidos a partir del espectro de Renji. De esta forma, se realiza un análisis clúster caracterizando cada una de las estaciones por un vector cuyas componentes son los valores de D_1 y D_2 (Tabla 4.16). Se obtienen así dos regiones, cada una de ellas con 19 y 34 estaciones respectivamente. Para la primera de las subregiones (Región 1 en la Tabla 4.18) solo una estación resultó discordante y se eliminó del análisis: estación de San Rafael con un valor de D_i de 3,45. El valor del estadístico H para la Región 1 fue de 0,33 (Tabla 4.18), por lo que esta región se considera homogénea. Para la región compuesta de los 34 lugares restantes, el valor de H obtenido fue de 1,89 después de eliminar las estaciones discordantes (Liguay, Los Queñes, Quella y Talca), por lo que esta región se considera discordante y ha de subdividirse en otras regiones.

Se realiza pues un nuevo análisis clúster con todas las estaciones que no se incluyeron en la Región 1 (Tabla 4.18), obteniéndose nuevamente dos grupos con 20 y 15 estaciones respectivamente. Para el primero de los grupos, las estaciones de Liguay ($D_i = 3,23$), San Rafael ($D_i = 3,55$) y Talca ($D_i = 3,03$), resultaron discordantes, mientras que el resto de las estaciones (17) forma una región homogénea, tal y como se deriva del valor del estadístico H obtenido: 0,76 (Región 2 en la Tabla 4.18).

Para la subregión formada por 15 estaciones no se encontró ninguna estación discordante y el valor obtenido para H fue de 1,00. Considerando que este último valor es el que marca la división entre homogeneidad y posible heterogeneidad, y teniendo en cuenta que hay una serie de estaciones que se han eliminado de las regiones homogéneas anteriores debido a su discordancia, es razonable formar una nueva subregión formada ahora por 18 estaciones. Esta nueva subregión cuenta con una estación discordante, San Rafael, con un valor de D_i de 3,18, que nuevamente se elimina del análisis. Las 17 estaciones restantes componen una región que resulta homogénea (Región 3 en Tabla 4.18) en base al valor obtenido del estadístico H (0,80).

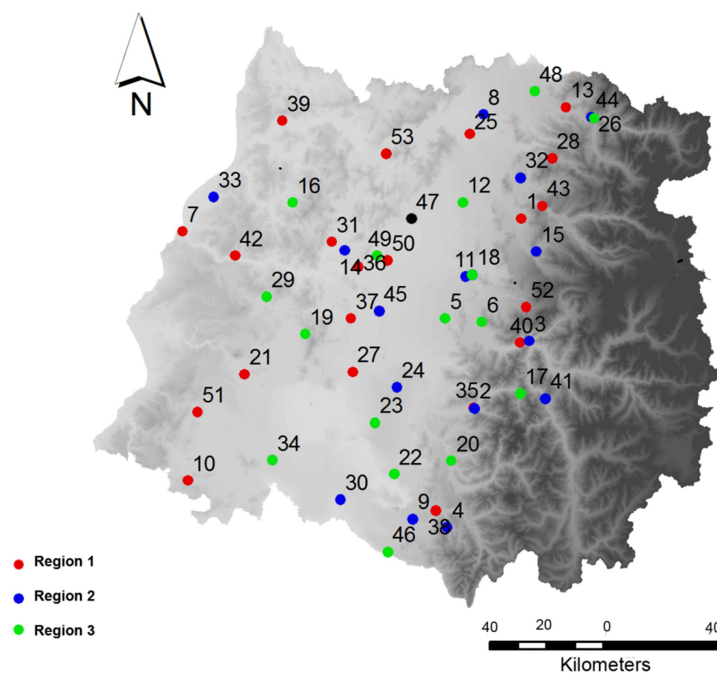


Figura 4.7. Regiones homogéneas obtenidas para la región del Maule (Chile).

La Figura 4.7 muestra las tres regiones homogéneas obtenidas tras los análisis realizados en tres colores diferentes: rojo para la Región 1, azul para la Región 2 y verde para la Región 3. Solo una estación (San Rafael) aparece en color negro ya que no ha sido posible incluirla en ninguna de las tres regiones. Para esta estación y con los

datos disponibles, solo sería posible el análisis local de frecuencias a la hora de analizar la precipitación máxima diaria anual.

Los resultados obtenidos pueden compararse con los obtenidos por García-Marín *et al.*, (2015) para la misma región. En esta última publicación hubo cinco estaciones que no pudieron incluirse en ninguna de las subregiones formadas. Por ello, la metodología de agrupación propuesta en esta Tesis basada en los valores de dimensión fractal de los datos disponibles, presenta una clara mejora en el proceso de agrupación en base a parámetros multifractales.

4.1.4.3. Conclusiones

A la vista de los resultados presentados en los apartados anteriores es posible agrupar series de datos de precipitación máxima diaria anual utilizando los valores de dimensión fractal obtenidos a partir del espectro de Rényi de las series de datos diarios. La novedad del método que aquí se presenta reside en que solo con dos valores que se obtienen con facilidad (D_1 y D_2) se obtienen regiones homogéneas al realizar un análisis regional de frecuencias de datos máximos anuales de lluvia diaria.

La metodología de análisis regional utilizada en este trabajo es la propuesta por Hosking y Wallis (1997). Mediante este análisis se ha comprobado que las 53 estaciones inicialmente disponibles con datos de precipitación diaria no formaban región homogénea desde el punto de vista de las precipitaciones máximas diarias anuales. Por ello, se han realizado varios análisis clúster basados en los valores de dimensión fractal obtenidos para cada una de las series disponibles. De esta manera, se han obtenido finalmente 3 regiones homogéneas dentro de la Región del Maule, quedando solo una estación fuera de cualquier análisis regional posible.

4.2. Identificación del tiempo mínimo entre eventos de precipitación mediante análisis multifractal

La precipitación es una variable de gran complejidad ampliamente utilizada en modelos hidrológicos de escorrentía, infiltración, erosión, intensidad-duración y frecuencia, etc. Teniendo en cuenta que la lluvia es un proceso intermitente, muchos estudios

hidrológicos y geomorfológicos emplean el concepto de evento de lluvia. Las propiedades de los eventos de precipitación dependen del criterio que se utilice para delimitarlos (e.g. Dunkerley, 2008). Por ello, hay que ser cuidadoso a la hora de definir un evento de lluvia en un determinado análisis ya que los distintos regímenes de escala pueden estar influenciados por la intermitencia de la lluvia e incluso por los períodos secos dentro de un determinado evento (e.g. Over y Gupta, 1996). Teniendo en cuenta que existe una gran variedad de definiciones de evento de lluvia (e.g. Dunkerley, 2008) hay que ser especialmente cauto a la hora de comparar los resultados obtenidos por los modelos que se basan en eventos de precipitación (Molini *et al.*, 2011).

Para definir un evento de lluvia es necesario fijar un período de tiempo sin lluvia, el Tiempo Mínimo entre Eventos (TME), necesario antes y después de cada evento (Dunkerley, 2008). El rango de TME que puede utilizarse es muy amplio y está directamente relacionado con la resolución temporal de los datos disponibles. El criterio a la hora de decidir el TME afectará a los resultados de los estudios posteriores, por lo que es importante analizar cuál es el TME más apropiado para un determinado lugar.

El objetivo de esta parte de Tesis Doctoral es identificar el TME más apropiado al trabajar con datos horarios de precipitación en dos localidades distantes, Málaga (España) y Bullileo (Chile) (véase apartado 3.1).

Existen trabajos que han puesto de manifiesto que la teoría de la Criticalidad Autoorganizada (del inglés ‘Self Organized Criticality’, SOC) puede utilizarse para describir la precipitación (e.g. Andrade *et al.*, 1998; Cowpertwait *et al.*, 2002; Peters y Christensen, 2002; Peters *et al.*, 2002, Telesca *et al.*, 2004; Arakawa, 2006; Bove *et al.*, 2006; Peters y Christensen, 2006; Peters y Neelin, 2006; Peters *et al.*, 2010; Deluca y Corral, 2014). La teoría SOC no analiza la secuencia temporal de eventos, sino su tamaño (Corral, 2015). Por lo tanto, la conexión entre la Criticalidad Autoorganizada y la precipitación debe analizarse a partir de la distribución del tamaño de los eventos de precipitación. A la hora de trabajar con la precipitación y la teoría SOC es habitual definir un evento de lluvia como una secuencia de registros consecutivos de lluvia mayores que cero y delimitados por registros de nula precipitación (Peters *et al.*, 2002). En esta Tesis se trabaja con TME de entre 1 y 6 horas sin precipitación para obtener los eventos de precipitación. La comparación entre las propiedades SOC de los eventos de lluvia para cada TME y las propiedades multifractales del conjunto de datos de lluvia,

permitirá identificar el TME más apropiado en cada una de las localidades objeto del estudio.

4.2.1. Definición y cálculo de eventos de lluvia

En esta Tesis se trabaja con la estructura de evento de lluvia propuesta por Peter y Christensen (2002). Según estos autores, un evento de lluvia puede definirse como una secuencia de registros de precipitación distintos de cero, $q(t)$, y su tamaño, M , se calcula según

$$M = \sum_i q(t) \Delta t \quad (4.18)$$

Donde Δt es igual a 1 hora en esta Tesis ya que se trabaja con datos de lluvia horaria. Los intervalos con registros nulos se consideran eventos secos (Figura 4.8). El número de eventos y su cantidad de lluvia puede variar en función del intervalo utilizado para separar un evento de otro. Si se usa el criterio propuesto por Peters y Christensen (2002) de manera estricta, cualquier intervalo sin lluvia marcaría el fin del evento en cuestión (Figura 4.8), estando los eventos de lluvia simplemente relacionados con la existencia o no de registro de precipitación. En esta Tesis se sigue el mismo criterio, pero añadiendo una nueva condición: la longitud mínima del intervalo temporal sin lluvia entre registros de precipitación (el citado TME), que variará entre 1 y 6 horas. De esta forma, en la Figura 4.8 se podrían identificar cuatro eventos si el TME es de una hora, o dos eventos para un TME de dos horas.

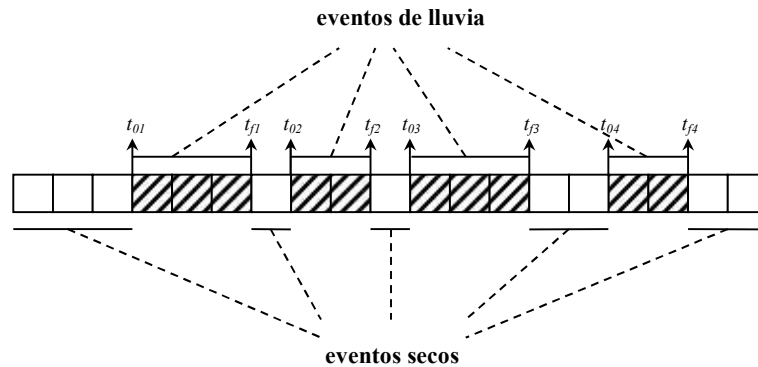


Figura 4.8. Definición de evento de lluvia

Como ya se ha mencionado, en esta parte de la Tesis se ha trabajado con datos horarios de precipitación en las localidades de Málaga y Bullileo, cuyos detalles se muestran en el apartado 3 de este documento. Aplicando la ecuación (4.18) y considerando valores de TME de entre 1 y 6 horas, se obtienen los eventos de lluvia para las citadas estaciones. La Tabla 4.19 recoge las principales características de estos eventos. Como puede observarse, el número de eventos disminuye al aumentar el valor del TME. Por el contrario, un aumento del TME produce un incremento de los valores medios de duración y altura de lluvia. Sin embargo, los valores de intensidad media permanecen constantes independientemente del valor del TME. Los valores de número de eventos, duración media y altura media de lluvia son más elevados para Bullileo, mientras que los valores de intensidad media de lluvia son ligeramente superiores para Málaga.

Tabla 4.19. Características de los eventos de lluvia para diferentes TME en Málaga y Bullileo.

	TME					
	1 h	2 h	3 h	4 h	5 h	6 h
Málaga						
Número de eventos	2642	2143	1883	1712	1610	1515
Duración media (h)	3,272	4,267	5,132	5,939	6,568	7,291
Altura media (mm)	5,200	6,411	7,296	8,025	8,533	9,068
Intensidad media (mm/h)	1,070	1,071	1,051	1,039	1,005	0,977
Intensidad máxima (mm/h)	14,25	13,00	12,59	12,59	12,59	12,56
Bullileo						
Número de eventos	6191	4935	4300	3903	3605	3399
Duración media (h)	6,059	7,856	9,312	10,564	11,768	12,784
Altura media (mm)	12,512	15,696	18,014	19,847	21,487	22,789
Intensidad media (mm/h)	0,950	0,936	0,933	0,931	0,929	0,923
Intensidad máxima (mm/h)	10,90	9,90	9,90	9,90	9,90	9,90

La Figura 4.9 muestra el diagrama de cajas y bigotes para las alturas y duraciones de eventos para diferentes TME en Málaga y Bullileo. El comportamiento es similar al ya descrito. Todos los valores tienden a aumentar al aumentar el TME, con los mayores para la estación de Bullileo. En cuanto a la duración del evento hay un aumento casi constante en la duración a medida que aumenta el TME. Sin embargo, los valores de altura de lluvia tienden a estabilizarse mostrando pequeñas diferencias para valores de TME superiores a 3 horas, siendo este comportamiento similar para ambas localidades.

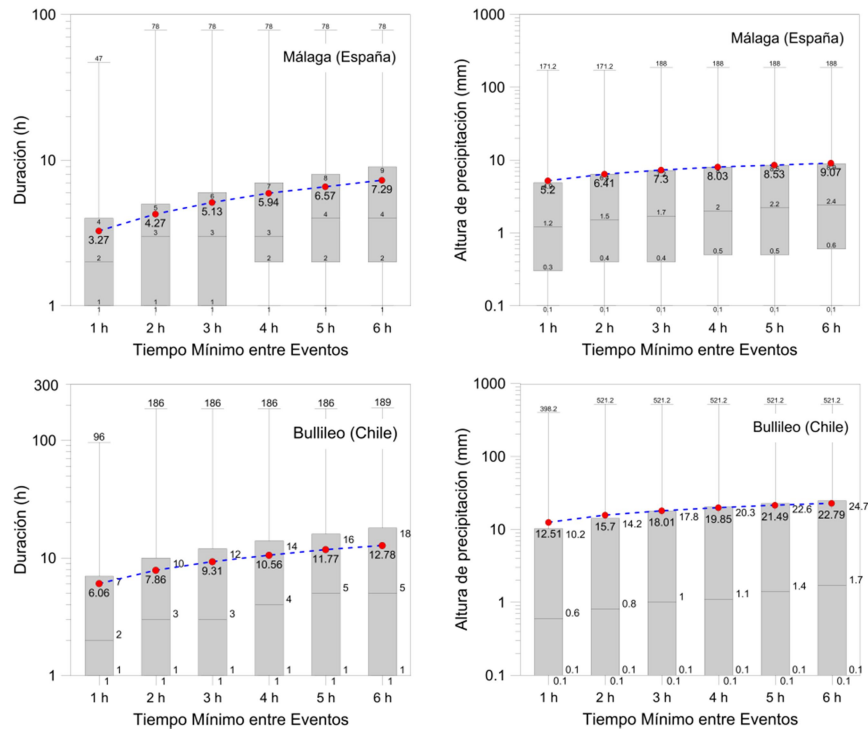


Figura 4.9. Diagrama de cajas y bigotes

4.2.2. Resultados de análisis multifractal y criticalidad autoorganizada

La Figura 4.10 muestra las funciones empíricas y teóricas escaladoras de momentos $K(q)$ obtenidas para Bullileo y Málaga a partir de los datos horarios de precipitación disponibles. Los parámetros de la función $K(q)$ teórica (ecuación 2.7) se han obtenido aplicando la técnica del momento de traza doble. Los valores obtenidos para el

parámetro α fueron de 0,640 para Bullileo y 0,533 para Málaga. Según estos valores, la lluvia en estas localidades pertenece a la clase universal, con $1 < \alpha < 0$ (Lovejoy y Schertzer, 1990). Resultados similares han sido documentados por diversos autores para otros lugares (Hubert *et al.*, 1993; Ladoy *et al.*, 1993; Tessier *et al.*, 1993; de Lima y Grasman, 1999; Garcia-Marin *et al.*, 2008a).

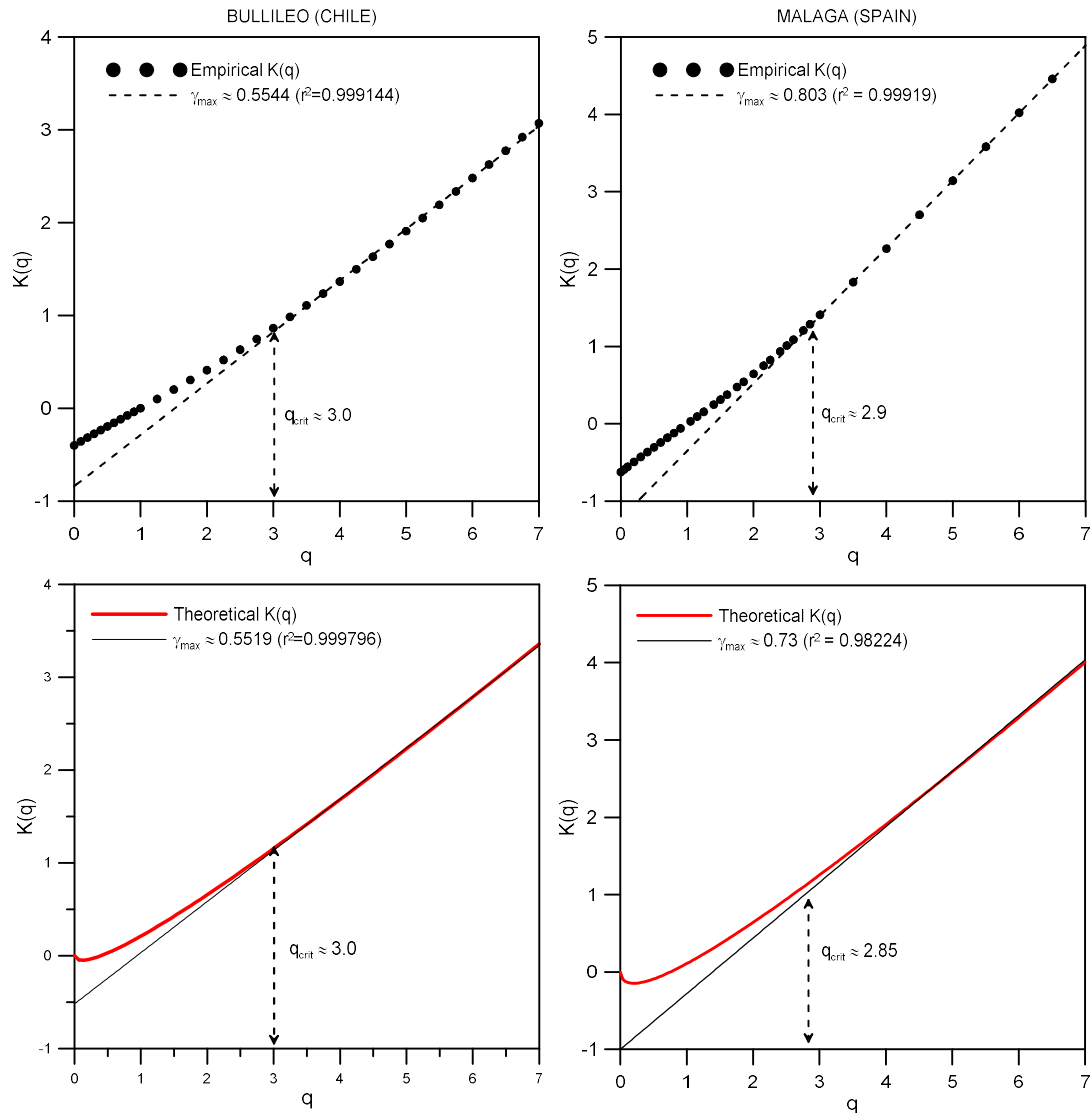


Figura 4.10. Funciones $K(q)$ empíricas y teóricas para Bullileo y Málaga.

Los valores de C_I fueron de 0,192 y 0,358 para Bullileo y Málaga, respectivamente. Los valores del exponente espectral necesarios para obtener el valor del parámetro H , han sido de 1,001 para Bullileo y 0,570 para Málaga. Teniendo en cuenta los valores de $K(2)$ de la función empírica escaladora de momentos (0,4104 para Bullileo y 0,644 para Málaga), los valores de H obtenidos han sido de 0,2108 para Bullileo y 0,107 para Málaga.

Todas las funciones $K(q)$ en la figura 4.10 tienen apariencia no lineal, indicando la estructura temporal multifractal del proceso de la lluvia para el rango de escala determinado a partir de los momentos estadísticos de los datos analizados. La función $K(q)$ ofrece información importante relativa al proceso de la precipitación, como ya se ha comentado en apartados anteriores de esta Tesis.

El valor de γ_{max} permite explicar cierta información relacionada con los eventos raros o extremos presentes en una serie temporal (e.g. Veneziano y Furcolo, 2002; de Lima y de Lima, 2009; García-Marín *et al.*, 2013b). El máximo valor corresponde a Málaga tanto en la función empírica (0,803) como en la teórica (0,730), lo que pone de manifiesto que las intensidades altas de lluvia destacan por los momentos más altos y sugiere un comportamiento más violento de la precipitación (e.g. de Lima y de Lima, 2009). Este comportamiento también puede observarse en la Tabla 4.19, donde las intensidades máximas de lluvia son mayores para Málaga (de 12,56 a 14,25 mm/h) que para Bullileo (9,90 a 10,90 mm/h) para todos los TME considerados.

Continuando con la Figura 4.10, los valores de $K(0)$ pueden relacionarse con los valores bajos de lluvia presentes en el conjunto de datos analizados, con los ‘ceros’ del proceso (de Lima, 1998). Para Málaga, $K(0)$ es -0,6400 mientras que para Bullileo este valor es de -0,4009. El porcentaje de horas sin lluvia en el conjunto de datos es de 99,68% para Málaga y 89,03% para Bullileo, por lo que el valor más bajo de $K(0)$ corresponde a la estación con el porcentaje más elevado de horas sin lluvia.

Los valores bajos de precipitación pueden también caracterizarse por $K(0) = -Cs$. Cs es la codimensión del proceso, siendo la dimensión fractal del conjunto de datos $D = 1 - Cs$ (e.g. García-Marín *et al.*, 2013b). La dimensión fractal D tiene valores próximos a 1 para

series ‘saturadas’ (con pocos valores de datos iguales a cero) (e.g. De Lima y De Lima, 2009). Los valores de dimensión fractal obtenidos para Bullileo y Málaga son de 0,5996 y 0,3600, respectivamente, lo que concuerda con el porcentaje de datos de no-lluvia en cada localidad (más elevado en Málaga que en Bullileo).

Tanto la función $K(q)$ teórica como la empírica muestran una zona lineal cuando los momentos q superan el valor crítico q_{crit} , siendo este último de 2,9 y 3,0 para Málaga y Bullileo respectivamente en la función empírica. El mismo valor se ha obtenido en Bullileo para la función $K(q)$ teórica, mientras que en Málaga el valor obtenido es ligeramente inferior (2,85). Teniendo en cuenta que ambos valores son muy similares, será el valor empírico el que servirá de base en los pasos posteriores.

Para poder encontrar relación entre estos dos últimos valores y la complejidad del fenómeno de la precipitación descrita por la teoría SOC, se ha obtenido el número de eventos de una determinada cantidad $N(x)$ para Málaga y Bullileo, para valores del TME que oscilan entre 1 y 6 horas. Tras representar $N(x)$ frente al tamaño de evento x (mm), se ha ajustado la función de distribución $\overline{N}(x)$ (ecuación 2.21) para los eventos entremos en Málaga (Figura 4.11) y Bullileo (Figura 4.12).

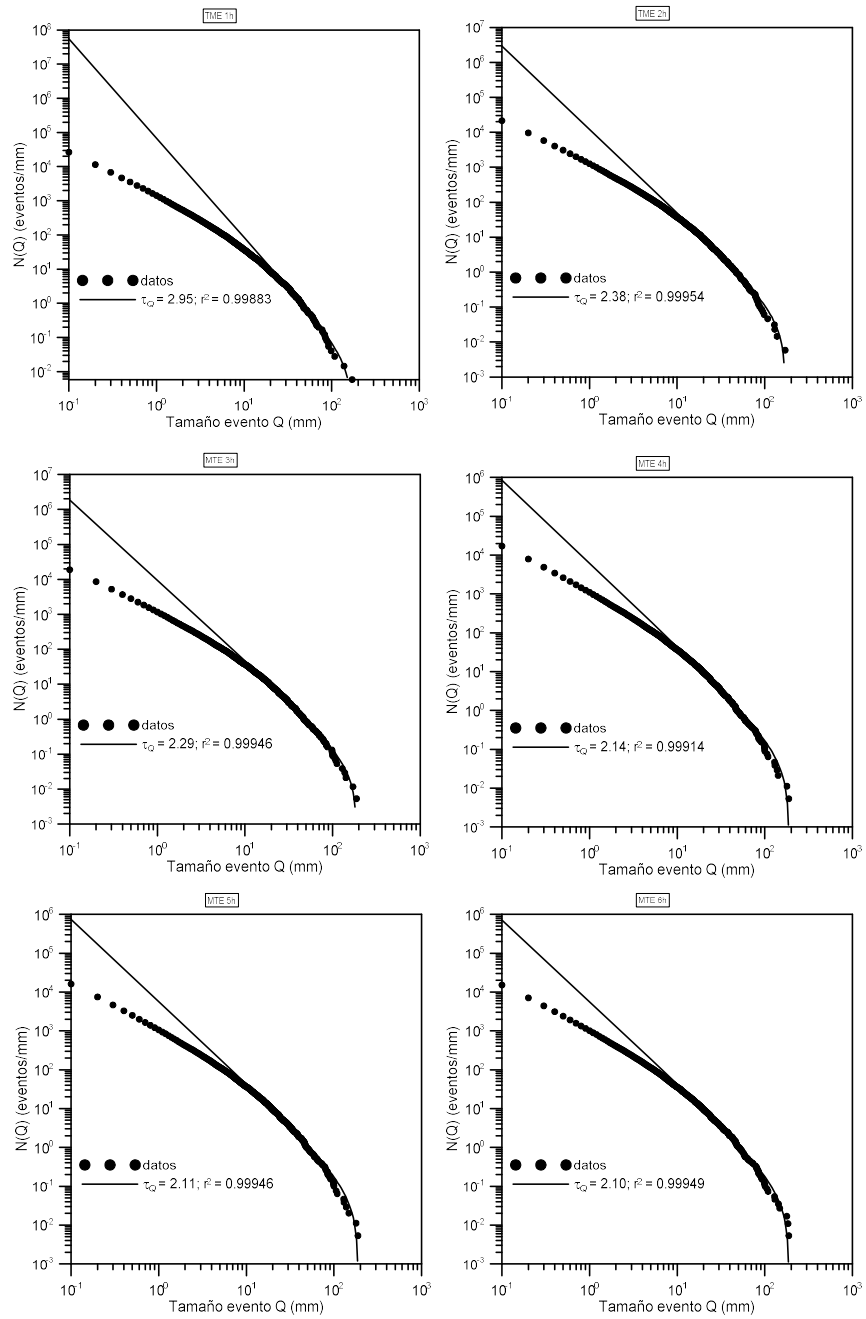


Figura 4. 11. Ajustes obtenidos con la función de distribución empírica para los tamaños de los eventos de lluvia en Málaga con MTE entre 1 y 6 horas.

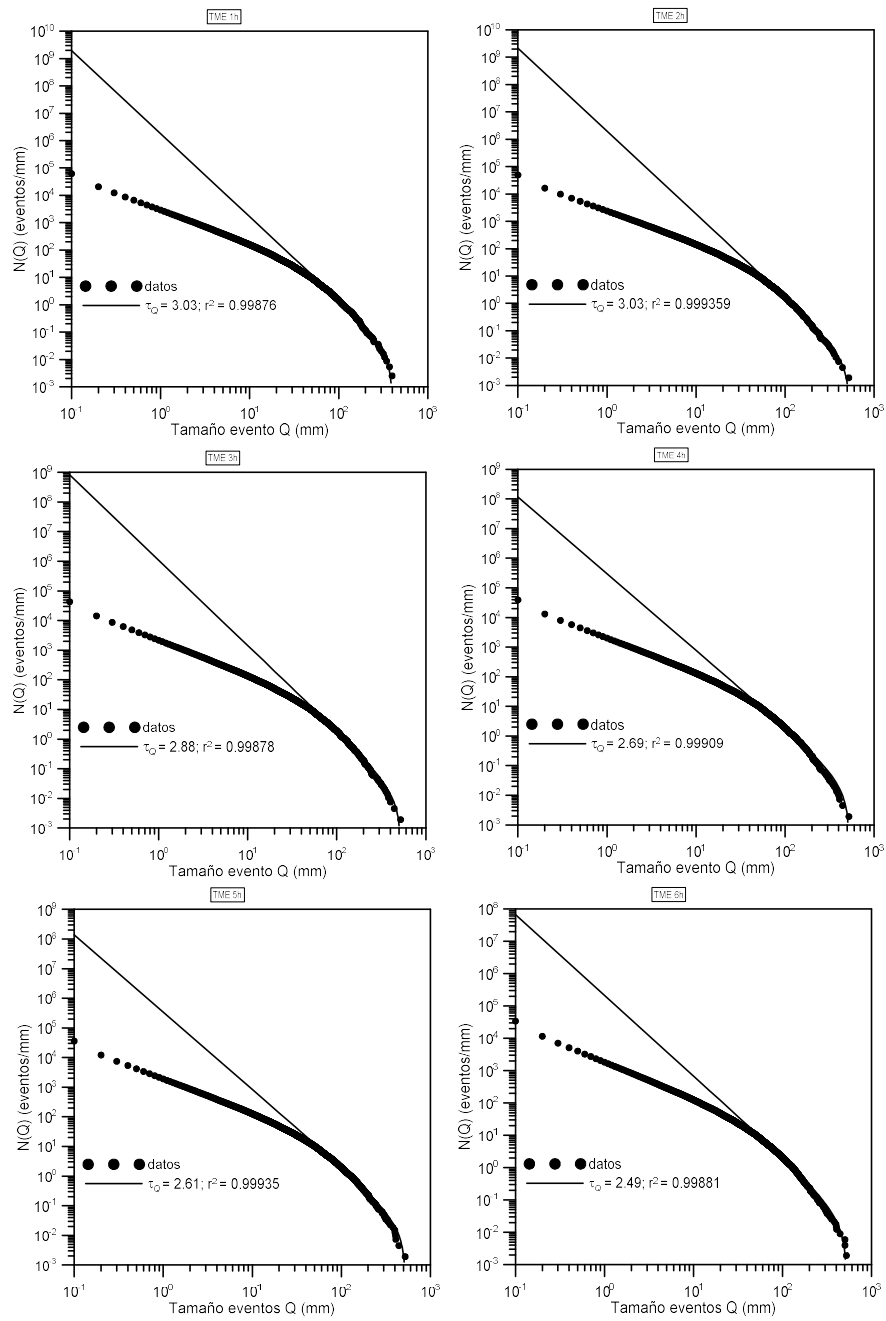


Figura 4.12. Ajustes obtenidos con la función de distribución empírica para los tamaños de los eventos de lluvia en Bullileo con MTE entre 1 y 6 horas

Como puede observarse en las figuras 4.11 y 4.12, el valor del exponente τ oscila entre 2,10 (TME 6 horas) y 2,91 (TME 1 hora) para Málaga, y entre 2,49 (TME 6 horas) y 3,03 (TME 1 y 2 horas) para Bullileo. Para ambas estaciones, el exponente τ disminuye a medida que el TME aumenta, obteniéndose además valores mayores para Bullileo que para Málaga.

Si se comparan los valores de q_{crit} para Málaga (2,9) y Bullileo (3,0) con el valor del exponente τ para los diferentes valores de TME considerados (Figuras 4.11 y 4.12), se observa que ambos valores son similares para TME de 1 hora en Málaga (2,91) y TME de 1 y 2 horas en Bullileo (3,03). La equivalencia de ambos valores es una de las consecuencias de la multifractalidad (e.g. Schertzer *et al.*, 1993; Schertzer y Lovejoy, 1994) que producen los sistemas críticamente autoorganizados (Bak *et al.*, 1987, 1988). Por lo tanto, teniendo en cuenta que ambos valores están relacionados y que dan información de la invarianza de escala presente en el proceso de la precipitación, el valor del TME más adecuado para una determinada localidad será aquel que proporcione una mayor similitud entre los valores de q_{crit} y τ . Esta similitud se obtiene para eventos separados por 1 hora sin lluvia en Málaga y para eventos separados por 1 y 2 horas en Bullileo.

4.2.3. Conclusiones

La estructura de los eventos de lluvia condiciona el comportamiento de diversos procesos a nivel de cuenca como la respuesta hidrológica, la ocurrencia de deslizamientos, el sellado del suelo, la erosividad, el lavado de componentes agroquímicos, entre otros. Para modelar estos y otros procesos, deben identificarse de manera correcta los eventos de lluvia a partir de los registros disponibles, estando las propiedades de dichos eventos (duración, altura, intensidad...) condicionadas por el valor del TME considerado.

En esta Tesis se presenta una nueva metodología para determinar el TME más apropiado para separar eventos de lluvia en un conjunto de datos con resolución horaria. Esta metodología combina dos teorías: el análisis multifractal y la teoría de la

criticalidad autoorganizada, ambas relacionadas con las propiedades de invarianza de escala del sistema que se analice.

Para probar la utilidad de la metodología propuesta, se han seleccionado dos estaciones distantes, una en España, Málaga, y otra en Chile, Bullileo. Para ambas estaciones, se dispuso de datos horarios de precipitación con una longitud de 30 años en Málaga y 39 años en Bullileo.

A partir de las funciones exponente escaladora de momentos empírica y teórica se obtuvo importante información relacionada con los eventos de lluvia extremos y con las horas sin lluvia. El mayor valor γ_{max} se obtuvo en Málaga, sugiriendo un comportamiento más violento de la lluvia en esta localidad que en Bullileo, tal y como realmente sucede. Los valores de $K(0)$ coinciden también con la presencia de horas secas en las series de datos analizadas, con un valor más bajo en Málaga al presentar un mayor porcentaje de horas sin lluvia. El valor de las horas sin lluvia también concuerda con los valores de la dimensión fractal (D), más cercanos a 1 para Bullileo que para Málaga.

Para el objetivo de esta parte de la tesis, el valor más importante de la función $K(q)$ es el valor del momento crítico $q_{crit.}$, ya que su coincidencia con el valor del exponente τ al ajustar la función de distribución $\overline{N}(x)$ a los eventos extremos obtenidos con TME de entre 1 y 6 horas, permite obtener el valor más apropiado del TME para cada estación, siendo 1 hora para Málaga y entre 1 y 2 horas para Bullileo.

En esta Tesis se ha trabajado con datos horarios de lluvia, por lo que, si la resolución temporal de la base de datos disponible cambia, habría que repetir el procedimiento ya que el tiempo entre eventos está directamente relacionado con la resolución temporal de los datos de lluvia.

5. CONCLUSIONES GENERALES Y LÍNEAS FUTURAS

La lluvia es un fenómeno complejo con una alta variabilidad espacial y temporal, y de su conocimiento detallado dependen los resultados de infinidad de cálculos hidrológicos e hidráulicos, estudios de cambio climático, predicciones futuras, etc. Existen diversos análisis que permiten explicar y describir el comportamiento de la precipitación, siendo el análisis multifractal con sus diferentes algoritmos uno de ellos. El análisis de invarianza de escala o análisis multifractal aporta información inherente al proceso o sistema que se esté analizando, información que cambia si cambia la dinámica del proceso. Cosa que no ocurre en los modelos que tratan de describir la lluvia desde un punto estadístico.

En esta Tesis doctoral se han utilizado las propiedades de invarianza de escala de la precipitación para dos propósitos principales. Por un lado, se ha estudiado su uso para agrupación de series de datos de lluvia máxima anual y por otro, para la separación de eventos de eventos de precipitación. En ambos casos se han manejado series de datos de localizaciones distintas y distantes.

En el caso de la agrupación de series de datos de precipitación máxima anual, y mediante la metodología de análisis regional de frecuencias propuesta por Hosking y Wallis (1997), se ha conseguido formar regiones de datos precipitación máxima diaria anual en base parámetros multifractales que describen las series históricas de lluvia diaria en las localidades disponibles. Se han realizado tres tipos de agrupaciones en base a estos parámetros, suponiendo cada una de ellas un avance respecto a la anterior. En la primera de ellas hubo que recurrir a la inclusión de otro parámetro no multifractal pero característico de cada serie de datos analizada. En el segundo análisis, simplemente añadiendo un nuevo parámetro multifractal, se pudieron agrupar las series de datos en regiones homogéneas sin necesidad de recurrir otras características de la lluvia. El mayor de los avances se ha conseguido con la tercera aproximación, donde la agrupación se ha conseguido a partir de dos nuevos parámetros multifractales de fácil obtención. La mejora conseguida en este tercer paso queda avalada por la comparación de los resultados del mismo con los obtenidos en la misma región de estudio, pero con otros parámetros multifractales publicado en García-Marín *et al.* 2015.

Por lo tanto, en base a los resultados obtenidos en esta Tesis doctoral puede afirmarse que la caracterización multifractal de datos de lluvia se puede usar de manera directa para identificar regiones homogéneas cuando se realiza un análisis regional de frecuencias, o como base de cualquier análisis estadístico para agrupar series de datos.

Continuando con la aplicación del comportamiento multifractal de la precipitación y gracias a la relación de este comportamiento con la criticalidad autoorganizada también existente en el fenómeno de la precipitación, se ha propuesto en esta Tesis un método que permite identificar el tiempo mínimo entre eventos de lluvia más adecuado a la hora de separar eventos de lluvia en una determinada localidad.

La metodología presentada en esta Tesis para decidir el tiempo entre eventos de lluvia más adecuado en un determinado lugar puede utilizarse siempre que sea necesario conocer información derivada de los eventos de lluvia, como en estudios de erosión, balance hídrico, interceptación, etc.

Como líneas futuras derivadas de la Tesis que aquí se presenta se pretende aplicar la metodología propuesta de agrupación en regiones a otras variables hidrometeorológicas, así como aplicar la metodología de separación de tiempo mínimo entre eventos de lluvia a series de datos de mayor resolución temporal con el objetivo de realizar estudios detallados de erosividad.

6. REFERENCIAS

- Andrade, R.F.S., Schellnhuber, H.J., Claussen, M., 1998. Analysis of rainfall records: possible relation to self-organized criticality. *Physica A* 254: 557-568.
- Arakawa, A., 2006. Atmospheric physics – Scaling tropical rain. *Nature Physics* 2, 373-374.
- Ariza-Villaverde, A.B., Jiménez-Hornero, F.J., Gutiérrez de Ravé, E., 2013. Multifractal analysis applied to the study of the accuracy of DEM-based stream derivation. *Geomorphology* 197, 85-95.
- Ayuso, J.L., 2012. Análisis Regional de Frecuencias de Valores Extremos. Departamento de Ingeniería Rural. Universidad de Córdoba.
- Bak, P., 1997. *How Nature Works: The Science of Self-Organized Criticality*. Springer Verlag: New York.
- Bak, P., Tang, C., Wiesenfeld, K., 1987. Self-organized criticality: an explanation of $1/f$ noise. *Physical Review Letters* 59 (4), 381-384.
- Bak, P., Tang, C., Wiesenfeld, K., 1988. Self-organized criticality. *Physical Review A* 38, 364-374.
- Beable, M.E., McKerchan, A. I., 1982. Regional flood estimation in New Zealand. Water and Soil Technical Publication 20, Ministry of Works and Development, Wellington, N. Z.
- Bishop, C. M., Hinton, G., 1995. *Redes neuronales para reconocimiento de patrones*. Clarendon Press.
- Bonell, M., Sumner, G., 1992. Autumn and winter daily precipitation areas in Wales, 1982-1983 to 1986-1987. *International Journal of Climatology* 12, 77-102.
- Bove, R., Pelino, V., De Leonibus, L., 2006. Complexity in rainfall phenomena. *Communications in Nonlinear Science and Numerical Simulation* 11(6), 663-776.
- Castro, J., Cârsteanu, A., Flores, C., 2004. Intensity-duration-area frequency functions for precipitation in a multifractal framework. *Physica A: Statistical Mechanics and its Applications* 338 (1-2), 206-210.

- Corral, A., 2015. Scaling in the timing of extreme events. *Chaos, Solitons & Fractals* 74, 99-112.
- Corral, A., Deluca, A., 2013. Fitting and goodness-of-t test of non-truncated and truncated power-law distributions. *Acta Geophysica* 61(6), 1351-1394.
- Cowpertwait, P.S.P., Kilsby, C.G., O'Connell, P.E., 2002. A space-time Neyman-Scott model of rainfall: empirical analysis of extremes. *Water Resources Research* 38, 1131. DOI: 10.1029/2001WR000709.
- Davis, A., Marshak, A., Wiscombe, W., Cahalan, R., 1994. Multifractal characterization of non stationarity and intermittency in geophysical fields: observed, retrieved or simulated. *Journal of Geophysical Resources* 99, 8055-8072.
- De Bartolo, S.G., Gabriele, S., Gaudio, R., 2000. Multifractal behavior of river networks. *Hydrology and Earth Sciences* 4(1), 105-112.
- De Lima, M.I.P., 1998. Multifractals and the Temporal structure of Rainfall. PhD. Thesis, Wageningen Agricultural University, 229 pp.
- De Lima, M.I.P., Grasman, J., 1999. Multifractal analysis of 15 min and daily rainfall from a semi-arid región in Portugal. *Journal of Hydrology* 220, 1-11.
- De Lima, M.I.P., De Lima, J.L.M.P., 2009. Investigating the multifractality of point precipitation in the Madeira archipelago. *Nonlinear Processes in Geophysics* 16, 299-311.
- Deluca, A., Corral, A., 2014. Scale invariant events and dry spells for medium-resolution local rain data. *Nonlinear Processes in Geophysics* 21, 555-567.
- Dunkerley, D.L., 2008. Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site. *Hydrological Processes* 22, 5024-5036.
- Easterling, D.R., 1989. Regionalization of thunderstorm rainfall in the contiguous United States. *International Journal of Climatology* 9, 567-579.
- Erro Eslava, J., 2011. Análisis Regional de las precipitaciones diarias extremas en la cuenca del río Arga. Trabajo Profesional Fin de Carrera. Escuela Técnica Superior de Ingenieros Agrónomos. Universidad Pública de Navarra.

- Estévez, J., 2008. Diseño de un sistema integrado para el control de calidad de datos de estaciones meteorológicas automáticas. Tesis Doctoral. Universidad de Córdoba.
- Estévez, J., Gavilán. P., García-Marín, A.P., 2010. Data validation procedures in agricultural meteorology. A prerequisite for their use. 10th Annual Meeting of the European Meteorological Society/8th European Conference on Applied Climatology, 13-17 Septiembre, Zurich, Suiza.
- Estévez, J., Gavilán. P., García-Marín, A.P., 2011a. Data validation procedures in agricultural meteorology. A prerequisite for their use. *Advances in Science and Research* 6, 141-146.
- Estévez, J., Gavilán. P., Giráldez, J.V., 2011b. Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology* 402, 144-154.
- Feder, J., 1988. *Fractals*, Springer, 283p.
- Feeny, B.F., 2000. Fast multifractal analysis by recursive box covering. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering* 10(9), 2277-2287.
- Feng, S., Hu, Q., Qian, Q., 2004. Quality control of daily meteorological data in China, 1951-2000: a new dataset. *International Journal of Climatology* 24, 853-870.
- Florez, L. and Bustaraca, C. 2008. *Fractales*. Computación gráfica.
- Fraedrich. K., Larnder, C., 1993. Scaling regimes of composite rainfall time series. *Tellus Series A – Dynamic Meteorology and Oceanography* 45A, 289-298.
- García Marín, A., 2007. Análisis multifractal de series de datos pluviométricos en Andalucía. Tesis Doctoral. Universidad de Córdoba.
- García-Marín, A.P., Jiménez-Hornero, F.J., Ayuso-Muñoz, J.L., 2007. Applying multifractality and the self-organized criticality theory to describe the temporal rainfall regimes in Andalusia (southern Spain). *Hydrological Processes* 22, 295–308
- García-Marín, A.P., Jiménez-Hornero, F.J., Ayuso-Muñoz, J.L., 2008. Universal multifractal description of an hourly rainfall time series from a location in southern Spain. *Atmósfera*, nº 21 (4), 347-355.

- García-Marín, A., Ayuso Muñoz, J.L., Taguas Ruiz, E.V., Estévez, J., 2011. Regional analysis of the annual maximum daily rainfall in the province of Málaga (Southern Spain) using the principal component analysis. *Water and Environment Journal* 25 (4), 522-531.
- García-Marín, A.P., Estévez, J., Jiménez-Hornero, F.J., Ayuso-Muñoz, J.L., 2013. Multifractal analysis of validated wind speed time series. *Chaos* 23, 013133.
- García-Marín, A.P., Ayuso-Muñoz, J.L., Jiménez-Hornero, F.J., Estévez, J., 2013a. Selecting the best IDF model by using the multifractal approach. *Hydrological Processes* 27, 433-443.
- García-Marín, A.P., Estévez, J., Sangüesa-Pool, C., Pizarro-Tapia, R., Ayuso-Muñoz, J.L., Jiménez-Hornero, F.J., 2015. The use of the exponent $K(q)$ function to delimit homogeneous regions in regional frequency analysis of extreme annual daily rainfall. *Hydrological Processes*, 29, 139-151.
- Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., Berne, A., Lovejoy, S., 2014. Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data. *Atmospheric Research* 138, 125-138.
- Grassberger, P., 1983. Generalized dimensions of strange attractors. *Physics Letter A* 97, 227-230.
- Grassberger, P., Procaccia, I., 1983. Measuring the strangeness of strange attractors. *Physica D* 9, 189-208.
- Greenwood, J., Landwehr, J., Matalas, N., Wallis, J., 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressed in inverse form. *Water Resources Research* 15(6), 1049-1054.
- Guadagnini, A., Neuman, S.P., Riva, M., 2012. Numerical investigation of apparent multifractality of samples from processes subordinated to truncated fBm. *Hydrological Processes* 26 (19), 2894-2908.
- Gupta, V.K., Waymire, E.C., 1993. A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology* 32, 251-267.
- Guttman, N.B., 1993. The use of L-moments in the determination of regional precipitation climates. *Journal of Climate* 6, 2309-2325.

- Halsey, T.C., Jensen, M.H., Kadanoff, L.P., Procaccia, I., Shraiman, B.I., 1986. Fractal measures and their singularities: the characterization of strange sets. *Physical Review A* 33, 1141-1151.
- Haylock, M., Nicholls, N., 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. *International Journal of Climatology* 20, 1533-1541.
- Hentschel, H.G.E., Procaccia, I., 1983. The infinite number of generalized dimension of fractals and strange attractors. *Physica D: Nonlinear Phenomena* 8, 435-444.
- Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrological Science Journal* 30, 85-109.
- Hosking, J.R.M., Wallis, J.R., 1988. The effect of intersite dependence on regional flood frequency analysis. *Water Resources Research* 24, 588-600.
- Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society* 1352, 105-124.
- Hosking, J.R.M., 1992. Moments or L-moments? An example comparing two measures of distributional shape. *The American Statistician* 46 (3), 186-189.
- Hosking, J.R.M., Wallis, J.R., 1997. Regional frequency analysis. Cambridge University Press. 224 pp.
- Hubbard, K.G., Goddard, S., Sorensen, W.D., Wells, N., Osugi, T.T., 2005. Performance of quality assurance procedures for an applied climate information system. *Journal of Atmospheric and Oceanic Technology* 22, 105-112.
- Hubert, P., Tessier, Y., Lovejoy, S., Schertzer, D., Ladoy, P., Carbonnel, J.P., Violette, S., 1993. Multifractals and Extreme Rainfall events. *Geophysical Research Letters* 20, 931-934.
- Jiménez Hornero, F.J., Pavón-Domínguez, P., Gutiérrez de Ravé, E., Ariza-Villaverde, A.B., 2011. Joint multifractal description of the relationship between wind patterns and land surface temperature. *Atmospheric Research* 99, 366-376.

- Jingyi, Z., Hall, M.J., 2004. Regional flood frequency analysis for Gang-Ming river basin in China. *Journal of Hydrology* 296, 98-117.
- Kite, G. W. 1977. Frequency and risk analysis in hidrology. Water Resources Publications. Fort Collins, Colorado.
- Kravchenko, A.N., Boast, C.W., Bullock, D.G., 1999. Multifractal analysis of soil spatial variability. *Agronomy Journal* 91, 1033-1041.
- Krzyszczak, J., Baranowski, P., Hoffmann, H., Zubik, M., Sławiński, C., 2017. Analysis of Climate Dynamics Across a European Transect Using a Multifractal Method. *Advances in Time Series Analysis and Forecasting*, DOI 10.1007/978-3-319-55789-2_8
- Ladoy, P., Schmitt, F., Schertzer, D., Lovejoy, S., 1993. The multifractal temporal variability of Nimes rainfall data. *ComptesRendus de l'Academie des Sciences Serie II* 317 (6), 775-782.
- Langousis, A., Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Research* 43W02422: DOI: 10.1029/2006WR005245.
- Langousis, A., Veneziano, D., Furcolo, P., Lepore, C., 2009. Multifractal rainfall extremes: theoretical analysis and practical estimation. *Chaos, Solitons and Fractals* 39, 1182-1194.
- Lavallée, S., Lovejoy, S., Schertzer, D., Ladoy, P., 1993. Nonlinear variability and Landscape topography: analysis and simulation. In *Fractals in Geography*, Lam N, De Cola L (Eds.). Prentice Hall: Englewood Cliffs, 158-192.
- Lettenmaier, D.P., Potter, K.W., 1985. Testing flood frequency estimation methods using a regional flood generating model. *Water Resources Research* 21 (12), 190-1914.
- Lovejoy, S., Schertzer, D., 1990. Multifractals, universality classes, satellite and radar measurements of clouds and rain. *Journal of Geophysical Research* 95 (D3), 2021-2034.
- Lovejoy, S., Schertzer, D., 1995. Multifractals and rain. In: Kundzewicz, Z.W. (Ed.), *New Uncertainty Concepts in Hydrology and Water Resources*. Cambridge University Press- UNESCO International Hydrology Series, New York, pp. 61-103.

- Maity, A.K., Pratihari, R., Mitra, A., Dey, S., Agrawal, V., Sanyal, S., Banerjee, A., Sengupta, R., Ghosh, D., 2015. Multifractal Detrended fluctuation analysis of alpha and theta EEG rhythms with musical stimuli. *Chaos, Solitons and Fractals* 81, 52-67.
- Mandelbrot, B., 1972. Possible refinement of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In: *Statistical Models and Turbulence. Lecture Notes in Physics*, vol. 12, pp. 333-351.
- Mandelbrot, B., 1982. *The fractal geometry of nature*. Freeman, San Francisco, 460 pp.
- Molini, L., Parodi, A., Rebora, N., Craig, G.C., 2011. Classifying severe rainfall events over Italy by hydrometeorological and dynamical criteria. *Quarterly. Journal of the Royal Meteorological Society* 137, 148-154.
- Moreno, F., Roldán, J., 1999. Regionalization of daily precipitation stochastic model parameters, application to the Guadalquivir Valley in Southern Spain. *Physics and Chemistry of the Earth* 24, 65-71.
- Over, T.M., Gupta, V.K., 1994. Statistical analysis of mesoscale rainfall: dependence of a random cascade generator on large scaling forcing. *Journal of Applied Meteorology* 33, 1526-1543.
- Parisi, G., Frisch, U., 1985. Fully developed turbulence and intermittency. In: Ghil, M., Benzi, R., Parisi, G., (Eds.), *Proceedings International School of Physics Enrico Fermi, Course LXXXVIII, Italian Physical Society. Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics*. North-Holland, Amsterdam.
- Over, T.M., Gupta, V.K., 1996. A space-time theory of mesoscale rainfall using random cascades. *Journal of Geophysical Research* 101, 26319-26331.
- Pandey, G., Lovejoy, S., Schertzer, D., 1998. Multifractal analysis of daily river flows including extremes for basins of five to two million square kilometres, one day to 75 years. *Journal of Hydrology* 208, 62-81.
- Peters, O., Christensen, K., 2002. Rain: relaxations in the sky. *Physical Review E* 66: 036120.
- Peters, O., Hertlein, C., Christensen, K., 2002. A complexity view of rainfall. *Physical Review Letters* 88, 018-701.

- Peters, O., Christensen, K., 2006. Rain viewed as relaxational events. *Journal of Hydrology* 328, 46-55.
- Peters, O., Neelin, J.D., 2006. Critical phenomena in atmospheric precipitation. *Nature Physics* 2, 393-396.
- Peters, O., Deluca, A., Corral, A., Neelin, J.D., Holloway, C.E., 2010. Universality of rain event size distributions. *Journal of Statistical Mechanics: Theory and Experiment* 11: P11030.
- Rodríguez, R., Casas, M.C., Redaño, A., 2013. Multifractal analysis of the rainfall time distribution on the metropolitan area of Barcelona (Spain). *Meteorology and Atmospheric Physics* 121 (3-4), 181-197.
- Royer, J-F., Biaou, A., Chauvin, F., Schertzer, D., Lovejoy, S., 2008. Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario. *Comptes Rendus- Geoscience* 340 (7), 431-440.
- Saa, A., Gasco, G., Grau, J.B., Anton, J.M., Tarquis, A.M., 2007. Comparison of gliding box and box counting methods in river network analysis. *Nonlinear Processes in Geophysics* 14, 603-613.
- Saenz De Ormijana, F., Hidalgo, F.J., Santa, A., 1991. Estimación de precipitaciones máximas mediante el método regional del índice de avenida. *Revista de Obras Públicas* 138, 9-22.
- Satyanarayana, P., Srinivas, V.V., 2011. Regionalization of precipitation in data sparse areas using large scale atmospheric variables- a fuzzy clustering approach. *Journal of Hydrology* 405, 462-473.
- Schertzer, D., Lovejoy, S., Lavallée, D., 1993. Generic multifractal phase transitions and self-organized criticality. In: *Perdang, J.M., Lejeune, A. (Eds.), Cellular Automata: Prospects in Astrophysical Applications*. World Scientific, 216-227.
- Schertzer, D., Lovejoy, S., 1987. Physical modelling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Journal of Geophysical Research: Atmosphere* 92, 9693-9714.

- Schertzer, D., Lovejoy, S., 1988. Multifractal simulations and analysis of rain and clouds by anisotropic scaling multiplicative processes. *Atmospheric Research* 21, 337-361.
- Schertzer, D., Lovejoy, S., 1994. Multifractal generation of self-organized criticality. In: Novak, M.M. (Ed.), *Fractals in the Natural and Applied Sciences*. Elsevier, North-Holland, pp. 325-339.
- Seager, R., Ting, M., Held, I., Kushnir, Y., Lu, J., Vecchi, G., Huang, H.P., Harnik, N., Leetmaa, A., Lau, N.C., Li, C., Véléz, J., Naik, N., 2007. Model projections of an imminent transition to a more arid climate in southwestern North America *Science* 316, 1181-1184.
- Shafer, M.G., 1990. Regional analysis of precipitation annual maximum in Washington State. *Water Resources Research* 26, 119-131.
- Shafer, M.A., Fiebrich, C.A., Arndt, D.S., Fredrickson, S.E., Hughes, T.W., 2000. Quality assurance procedures in the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology* 17, 474-494.
- Sivakumar, B., 2001. Is a chaotic multifractal approach for rainfall possible? *Hydrological Processes* 15, 943-955 (with a minor correction in *Hydrological Processes* 15: 2381-2382).
- Srinivas, V.V., Tripathi, S., Rao, A.R., Govindaraju, R.S., 2008. Regional flood frequency analysis by combining self-organizing feature maps and fuzzy clustering. *Journal of Hydrology* 348, 148-166.
- Svensson, C., Olsson, J., Berndtsson, R., 1996. Multifractal properties of daily rainfall in two different climates. *Water Resources Research*. 32, 2463-2472.
- Telesca, L., Colangelo, G., Lapenna, V., Macchiato, M., 2004. On the scaling behavior of rain event sequence recorded in Basilicata region (southern Italy). *Journal of Hydrology* 296, 234-240.
- Telesca, L., Lovallo, M., 2011. Analysis of time dynamics in Wind records by means of multifractal detrended fluctuation analysis and Fisher-Shannon information plane. *Journal of Statistical Mechanics: Theory and Experiment* P0700. DOI: 10.1088/1742-5468/2011/07/P07001.

- Tessier, Y., Lovejoy, S., Schertzer, D., 1993. Universal multifractals in rain and clouds: theory and observations. *Journal of Applied Meteorology* 32, 223-250.
- Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., Pecknold, S., 1996. Multifractal analysis and modelling of rainfall and river flows and scaling, causal transfer functions. *Journal of Geophysical Research-Atmospheres* 101, 26427-26440.
- Valencia, J.L., Requejo, A.S., Gascó, J.M., Tarquis, A.M., 2010. A universal multifractal description applied to precipitation patterns of the Ebro River Basin. *Spain Climate Research* 44 (1), 17-25.
- Veneziano, D., Furcolo, P., 2002. Multifractality of rainfall and scaling of intensity-duration-frequency curves. *Water Resources Research* 38 (12), DOI: 10.1029/2001WR000372.
- Veneziano, D., Iacobellis, V., 2002. Multiscaling pulse representation of temporal rainfall. *Water Resources Research* 38, 11-38, DOI: 10.1029/2001WR000522.
- Veneziano, D., Langousis, A., Furcolo, P., 2006. Multifractality and rainfall extremes: a review. *Water Resources Research* 42, W06D15, doi:10.1029/2005WR004716.
- Yürekli, K., Modarres, R., 2007. Regionalization of maximum daily rainfall data over Tokat province, Turkey. *International Journal of Natural and Engineering Sciences* 1 (2), 1-7.

7. PUBLICACIONES DERIVADAS DE LA TESIS DOCTORAL

Artículo 1

Journal of Hydrology 529 (2015) 106–119



Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Delimiting homogeneous regions using the multifractal properties of validated rainfall data series



A.P. García-Marín^{*}, J. Estévez, M.T. Medina-Cobo, J.L. Ayuso-Muñoz

Dept. of Rural Engineering, University of Córdoba, P.O. Box 3048, 14080 Córdoba, Spain

ARTICLE INFO

Article history:

Received 13 March 2013

Received in revised form 10 April 2015

Accepted 16 July 2015

Available online 22 July 2015

This manuscript was handled by Geoff Syme, Editor-in-Chief, with the assistance of Alin Andrei Carsteanu, Associate Editor

Keywords:

Regional frequency analysis

Meteorological data validation

Homogeneous regions

Multifractal properties

SUMMARY

In this work, a regional frequency analysis of extreme annual rainfall data in Malaga (Southern Spain) has been performed. Rainfall records have been validated, applying various quality control tests as a pre-requisite before their use, ensuring their reliability and discarding anomalous data. For grouping the stations into potential homogeneous regions, the multifractal properties of daily rainfall data series recorded at 72 locations have been studied. The scaling of the rainfall moments has been analyzed and the empirical moments scaling exponent functions have been obtained. The corresponding multifractal values have been used to group stations into regions, resulting some of them homogeneous.

© 2015 Elsevier B.V. All rights reserved.

Artículo 2

HYDROLOGICAL PROCESSES
Hydrol. Process. **30**, 3507–3517 (2016)
Published online 12 June 2016 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10875

The identification of an appropriate Minimum Inter-event Time (MIT) based on multifractal characterization of rainfall data series

M. T. Medina-Cobo, A.P. García-Marín,* J. Estévez and J.L. Ayuso-Muñoz

Dept. of Rural Engineering, Leonardo Da Vinci Building, University of Córdoba, 14071 Córdoba, Spain

Abstract:

Different hydrological models of sediment entrainment, infiltration, overland flow production, soil loss, landslide occurrence, among others, need rainfall events data. The Minimum Inter-event Time (MIT) used to separate rainfall events affects their properties. Therefore, it is highly important their correct definition. Various event definitions are commonly used in hydrology, and a high variety of MIT is applied. In this work, an easy method is proposed for deciding the proper MIT in a certain place when dealing with hourly rainfall data series. This new methodology is based on the scale invariance properties of rainfall, analysed with both multifractal and Self-Organized Criticality theories. The similarity between some parameters from both analyses allows picking out the more appropriate MIT for the hourly rainfall data sets analysed, in Málaga (Southern Spain) and Bullileo (Central Chile). The method can be applied for different rainfall time resolutions according to the purpose of the model that needs the rain event information. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS rainfall event; multifractality; self-organized criticality; power-law

Received 9 February 2016; Accepted 6 April 2016

Artículo 3

Water Resour Manage (2017) 31:2333–2348
DOI 10.1007/s11269-017-1653-2



Obtaining Homogeneous Regions by Determining the Generalized Fractal Dimensions of Validated Daily Rainfall Data Sets

M. T. Medina-Cobo¹ · A. P. García-Marín¹ · J. Estévez¹ ·
F. J. Jiménez-Hornero² · J. L. Ayuso-Muñoz¹

Received: 20 April 2016 / Accepted: 29 March 2017 /
Published online: 5 April 2017
© Springer Science+Business Media Dordrecht 2017

Abstract Extreme rainfall data are widely used in several hydrological models and civil engineering design. Despite high temporal resolution rainfall data are not commonly available, daily rainfall data series are easily found. When these available data series are short in length the Regional Frequency Analysis (RFA) is a good tool to enlarge them by joining stations into homogeneous regions. This is by far, the most complicated step in RFA. This work presents a new method to form homogeneous regions of extreme annual daily rainfall data series. Daily rainfall data series from 53 weather stations in the Maule Region (Chile) have been used. Their fractal dimensions spectra have been obtained by applying the box counting method. Each station has been characterized by the fractal dimensions D1 and D2. A cluster analysis has been carried out based on these at-site characteristics and three regions have been obtained. After performing a RFA of extreme daily annual rainfall data series within each region they have shown as homogeneous. Only one of the available stations has not been possible to be included into any homogeneous regions, being the local frequency analysis the only suitable method to be applied at this location.

Keywords Regional frequency analysis · Homogeneous regions · Fractal dimensions